DISSERTATIONES SCHOLA DOCTORALIS SCIENTIAE CIRCUMIECTALIS, ALIMENTARIAE, BIOLOGICAE. UNIVERSITATIS HELSINKIENSIS

SATU RUOKOLAINEN

THE MADS WORLD OF FLORAL REGULATORS IN *GERBERA HYBRIDA*

DEPARTMENT OF AGRICULTURAL SCIENCES
FACULTY OF AGRICULTURE AND FORESTRY AND
FACULTY OF BIOLOGICAL AND ENVIRONMENTAL SCIENCES
DOCTORAL PROGRAMME IN PLANT SCIENCES
UNIVERSITY OF HELSINKI

Dissertationes Schola Doctoralis Scientiae Circumiectalis, Alimentariae, Biologicae

Universitas Helsinkiensis
Doctoral Programme in Plant Sciences (DPPS)
36/2020

THE MADS WORLD OF FLORAL REGULATORS IN *GERBERA HYBRIDA*

Satu Ruokolainen

Department of Agricultural Sciences
Faculty of Agriculture and Forestry
and
Faculty of Biological and Environmental Sciences

University of Helsinki
Finland

DOCTORAL DISSERTATION
To be presented, with the permission of the Faculty of Biological and Environmental Sciences of the University of Helsinki, for public discussion, in Auditorium 108, Forest Sciences Building, (Latokartanonkaari 7, Helsinki) on the October 22nd, 2020, at 12 o'clock noon.
Helsinki 2020

Supervised by Professor Teemu Teeri
Department of Agricultural Sciences
University of Helsinki

Professor Paula Elomaa
Department of Agricultural Sciences
University of Helsinki

Reviewed by Professor Hely Häggman
Department of Biology
University of Oulu

Dr. Annakaisa Elo
HAMK Bio Research Unit
Häme University of Applied Sciences (HAMK)

Opponent Dr. Michiel Vandenbussche
Le Centre National de la Recherche Scientifique
École Normale Supérieure de Lyon /Université de Lyon

Custos Professor Kurt Fagerstedt
Faculty of Biological and Environmental Sciences
University of Helsinki

The Faculty of Biological and Environmental Sciences uses the Urkund system (plagiarism recognition) to examine all doctoral dissertations.

Doctoral Programme in Plant Science (DPPD)

Dissertationes Schola Doctoralis Scientiae Circumiectalis, Alimentariae, Biologicae
ISSN 2342-5423 (print)
ISSN 2342-5431 (Online)

ISBN 978-951-51-6687-6 (print)
ISBN 978-951-51-6688-3 (PDF)

Hansaprint Oy
Vantaa 2020

Cover art by Heli Ruokolainen 2010

‘Melius sero quam numquam’

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I

Ruokolainen Satu, Ng, Yan Peng, Broholm, Suvi K., Albert, Victor A., Elomaa, Paula and Teeri, Teemu H. (2010) Characterization of *SQUAMOSA*-like genes in *Gerbera hybrida*, including one involved in reproductive transition. BMC Plant Biology 10: 128.

II

Ruokolainen, Satu, Ng, Yan Peng, Albert, Victor A., Elomaa, Paula and Teeri, Teemu H. (2010) Large scale interaction analysis predicts that the *Gerbera hybrida* floral E function is provided both by general and specialized proteins. BMC Plant Biology 10:129.

III

Ruokolainen, S., Ng, Yan Peng, Albert, Victor A., Elomaa, Paula and Teeri, Teemu H. (2011) Over-expression of the *Gerbera hybrida At-SOC1-like* gene *Gh-SOC1* leads to floral organ identity deterioration. Annals of Botany 107(9): 1491-1499.

The publications are referred to in the text by their roman numerals.

AUTHORSHIP STATEMENT

Publication I

SR designed the experiment and cloned *GSQUA2*, *GSQUA3*, *GSQUA4* and further characterized them as well as *GSQUA5* and *GSQUA6* by expression analyses. Transformation of *GSQUA2* was partially done by SR**.** YPN participated in expression analyses of *GSQUA5* and *GSQUA6* and cloned the full-length *GSQUA6*. SKB characterized the *GSQUA2* overexpression phenotype inflorescence by light stereo microscopy and SEM. VAA performed the phylogenetic analysis. The manuscript was drafted by SR and later revised together with VAA, PE, and THT

Publication II

SR designed the experiment together with PE and THT. SR performed the experiments together with technician Eija Takala. YPN participated in the early phase in setting up the Gateway system and cloning the necessary plasmids. VAA constructed the phylogenetic tree. The manuscript was drafted by SR and later revised together with VAA, PE and THT.

Publication III

SR performed the expression analyses, protein-protein interaction experiments, anthocyanin analysis and drafted the manuscript. YPN analysed the transformant lines by light microscopy and SEM and measured the cell lengths. VAA did the phylogenetic analysis. The manuscript was drafted by SR and revised together with VAA, PE and THT.

ABBREVIATIONS

cDNA	Complementary DNA
EST	Expressed sequence tag
FM	Floral meristem
GA	Gibberellic acid
mRNA	Messenger RNA
N/A	Not available
PCR	Polymerase chain reaction
PPI	Protein-protein interaction
RT-PCR	Reverse transcriptase polymerase chain reaction
SEM	Scanning electron microscopy
WGD	Whole genome duplication
Y2H	Yeast two hybrid
Y3H	Yeast three hybrid

ABSTRACT

The flowering process of plants is of great importance - both for the sexual reproduction of plants and human nutrition. Floral diversity has fascinated scientists for centuries, but the first important steps towards explaining the molecular puzzle of flowering were taken only thirty years ago when plant MADS box genes were discovered, and the classical ABC model was proposed. The focus of research has shifted from studying single genes to working on whole networks of genes and proteins - genomics and proteomics. The genomics of Arabidopsis (*Arabidopsis thaliana*), the well-known model plant, is advanced, but phenomena present in Arabidopsis might be lacking in other species, and *vice versa*. The model plant, Gerbera (*Gerbera hybrida*), belongs to the large Asteraceae (sunflower) family. While Gerbera flowering process shares several features with Arabidopsis and other model plants, it also shows great specialization. Gerbera bears a complex inflorescence containing hundreds of flowers of three different types, which differ in morphology and sex. The major aim of this thesis was to characterize Gerbera floral developmental genes, with special interest in genes related to A and E functions.

The ABC model and the extended ABCDE model are applicable to Gerbera. It has been previously shown that the Gerbera B and C function genes behave as the model predicts. The E function genes affect the development of the whole flower in Arabidopsis, being redundantly active in all four floral whorls. However, the Gerbera E function genes, *GRCD1* and *GRCD2*, are non-redundant and specialized in their tasks. Based on *GRCD4* and *GRCD5* expression patterns and PPI data presented in this thesis we proposed them to provide general E function in Gerbera. However, later this hypothesis was shown not to be completely accurate by RNAi transgenic lines that showed *GRC4* and *GCRD5* to be involved in Gerbera petal development.

The ABCDE model proposes that A function genes determine the developmental fate of sepals and petals. This function, however, is the most problematic of the model since a true A function seems only to be present in Arabidopsis. Majority of homologous genes from the other model plants execute only partial functions of *AP1.* Gerbera contains several genes related to *AP1* and its homologues *CAL* and *FUL,* but none of them supply the A function in the sense of the ABC model. All Gerbera *AP1*- and *FUL*-like genes GSQUAs display wide expression patterns, some of them present in all floral organs. Different *GSQUA genes* were transformed into Gerbera, but only *GSQUA2* overexpression lines produced a recurrent phenotype that was an early flowering, dwarf Gerbera. The function of *GSQUA2* was shown to be linked to floral transition.

Phylogenetic analysis showed *Gh-SOC1* to be distantly paralogous to Arabidopsis *SOC1*. In contrast to Arabidopsis SOC1, *Gh-SOC1* was expressed only in the floral parts of Gerbera and it did not promote flowering but altered inflorescence identity towards vegetative characteristics. Temporal expression pattern late in floral development and floral abundancy suggested *Gh-SOC1* to have a role in the late stages of Gerbera floral organ development.

The results presented in this thesis add to our understanding of inflorescence and floral development of Gerbera. The ABCDE model is applicable to Gerbera for B and C function, A function does not seem exist in Gerbera and E function is differentiated from Arabidopsis general E function. Despite close sequence similarity to Arabidopsis *SOC1*, *Gh-SOC1* function in Gerbera is related to floral development.

1 INTRODUCTION

The developing flower has fascinated scientists for centuries. In the 18th century, the Swedish botanist Carl von Linné built his classification of plants based on their floral and fruit properties. Linné based his hierarchical system on the number of stamens and their arrangement within the flower. Later during the same century, in 1790, the German poet Goethe expanded his area of expertise by publishing an essay on plant metamorphosis (Goethe, 1790). He based his proposal on the variation in the normal plant growth and the abnormalities in floral structures that can be found in nature. According to Goethe, flowers were modified leaves. Goethe's simple idea has appealed plant scientists for over 200 years, and it has turned out to be true. Leaves can be converted into petals with the expression of A, B, and E class genes in Arabidopsis (*Arabidopsis thaliana*) (Pelaz *et al.* 2001b). Without the presence of the SEP proteins (encoded by D class MADS box genes), the Arabidopsis floral organs lose their identity and start to resemble leaves (Ditta *et al.* 2004). Based on these and numerous other studies, flowers are considered as modified shoots bearing modified leaves (sepals, petals, stamens and carpel). The floral shoot is condensed to form the receptacle of the flower.

No wonder the subject has intrigued researchers and laymen equally for centuries as flowering has crucial importance for plant sexual reproduction. Flowers attract pollinators and make the pollination take place in the insect-and bird- pollinated plants. The outcome of plant sexual reproduction is essential to humans since fruits, vegetables and seeds feed the majority of people on earth. Understanding plant sexual reproduction and flower development is vital also from this perspective. Floral development has been extensively studied for decades and in the early 1990's a combinatorial model explaining the function of organ identity genes was proposed as the basis for floral development (Coen & Meyerowitz 1991).

1.1 The classical ABC model of floral development

Based on homeotic mutants, very similar to what Goethe observed at his time, a simple model explaining the abnormalities in the floral structure was proposed by Coen and Meyerowitz in 1991. This classical model consisted originally of three different functions called the A, B and C. In the floral whorl one, the action of A function gives rise to sepals. In the second whorl, functions A and B together define petals. Moving towards the center of the flower, B and C functions together define stamens, and carpel identity is controlled by C function alone. The model further requires that A and C functions are mutually antagonistic. If either one of them is absent, the other function is able to spread to all floral whorls. Later this model was extended by adding D and E functions. D function defines ovule development independently from carpel development, and acts later than A, B, and C

functions. E function is necessary for development of all floral organs (Angenent *et al*. 1995, Colombo *et al*. 1995, Theissen *et al.* 2001, Theissen & Saedler 2001, Ditta *et al*. 2004, Ma *et al*. 2005, Zahn *et al*. 2005).

Even though the classical ABC model explains the identity of floral organs as the functions of different homeotic genes (Coen & Meyerowitz 1991), it is a simplified model. Some authors are inclined to present the ABC model as (A)BC model, as the existence of the entire A function might be Arabidopsis specific (Litt & Irish 2003, Davies *et al.* 2006, Litt 2007, Pabón-Mora *et al.* 2012). Sepal identity might not be controlled by a conserved A function, but rather represent an independently evolved genetic program, which differentiates bracts from petals (Buzgo *et al.* 2004). The A function activity might indirectly lead to sepal and petal development by specifying floral meristem and repressing the C function in whorls one and two (Wollmann *et al*. 2010).

The ABC model explains reasonably well floral structures in core eudicots but extending the model to the monocots and many basal angiosperms adds a next level of complexity. Several modified models to cover these plants have been suggested based on the floral morphology and/or the patterns of gene expression (van Tunen *et al.* 1993, Bowman 1997, Kanno *et al*. 2003, Kramer *et al.* 2003, Buzgo *et al*. 2004, Kim *et al.* 2005). Some of the modified models for floral organ development will be discussed together with B function genes.

1.2 The structure and the molecular function of MADS box genes

Majority of genes responsible for A, B, C, D, and E functions encode MADS domain transcription factors (reviewed in e.g. Krizek and Flecther 2005). These genes are conserved within the plant and animal kingdoms and participate in homeotic functions. In addition to their well described role in floral development, the role of MADS box genes in various plant stress responses has started to emerge (reviewed in Castelán-Muñoz *et al*. 2019). MADS box genes were named after the first characterized representatives in each species. M stands for *MCM1*, the first MADS domain gene isolated from baker's yeast *Saccharomyces cerevisiae* (Passmore *et al.* 1988). Letter A derives from *AGAMOUS*, the C function gene isolated from Arabidopsis (Yanofsky *et al.* 1990), and D stands for the B function gene *DEFICIENS,* which was characterized from Antirrhinum (*Anthirrinum majus*) (Sommer *et al*. 1990, Schwarz-Sommer *et al*. 1992). *SRF*, serum response factor from humans finishes the word MADS (Norman *et al.* 1988). MADS box genes are divided in two classes (I and II) based on their structure (Alvarez-Buylla *et al.* 2000). Both types are found in animals, plants and fungi and they have most likely originated by duplication before the divergence of these kingdoms (De Bodt *et al.* 2003a). In comparison to other organisms, the MADS box gene family in plants is large. Arabidopsis, which has a small genome, has 107 MADS box genes (Pařenicová *et al.* 2003). The high number of

genes involved in transcriptional regulation has been suggested to originate from large-scale duplication events and is considered to increase phenotypic variation (evolution of novel morphologies), and thus increase biological complexity (De Bodt *et al.* 2003b).Typically, animal and fungi genomes only contain two to six MADS box genes (Gramzow & Theissen 2013).

Specific functions have been assigned to the different domains of these transcription factors. The MADS domain (58 amino acids, located at N terminus) encodes DNA binding, nuclear localization, and dimerization functions. Protein-protein interactions are mediated by helix containing K domain and C terminal region (Riechmann & Meyerowitz 1997, Theissen *et al*. 2000, Alvarez-Buylla *et al.* 2000). I (intervening) region and K domain are suggested to be involved in specific dimerization. C terminal region is necessary for formation of ternary complexes (Egea-Cortines *et al.* 1999, Honma & Goto 2001). It also encodes transcriptional activation function (Moon *et al.* 1999, Honma & Goto 2001) and harbours conserved motifs (Litt & Irish 2003, Vandenbussche *et al*. 2003b). Structure of MADS domain protein is shown in figure 1.

MADS	58 aa, highly conserved, DNA binding, dimerization, nuclear localization, sub domains α, β
I	27-42 aa, considerable sequence variability
K	70 aa, moderately conserved, keratin related, PPI, contains 3 sub domains K1, K2, K3
C	variable sequence, conserved motifs
N	appr. 15 aa, C function MADS domain protein N terminal extension

Figure 1. Structure of MADS domain protein. MADS and K box are the most conserved domains. Intervening region I and C terminus display more variability in sequence. C terminus harbors conserved protein motifs. N terminal extension before MADS domain is found in C function proteins.

Based on their structure, MADS box genes can be divided into types I and II. The majority of characterized MADS box genes belong to the type II, also called the MIKC MADS box genes (Műnster *et al.* 1997). In addition to the highly conserved MADS domain (M), these genes contain also a weakly conserved intervening sequence (I), a conserved K domain (keratin like, K) and a variable C terminal region (C) (Alvarez-Buylla *et al*. 2000). The MIKC-type MADS box genes can be further divided to MIKC^{C} and MIKC* type genes based on their exon-intron structure (Henschel *et al.* 2002). The split of the different MIKC genes happened in the ancestor of all land plants (Gramzow & Theissen 2010). The MIKC^{C} genes have been identified in angiosperms, gymnosperms, ferns, and mosses (Becker & Theissen 2003), whereas the MIKC* type was originally found in clubmosses and

mosses (Svensson *et al*. 2000, Henschel *et al.* 2002). Later MIKC* genes were discovered also in Arabidopsis (Kofuji *et al.* 2003). The function of MIKC* type genes is poorly understood (Nam *et al.* 2004). Most of the well-studied MADS box genes belong to the MIKC^{C} class.

Type II MADS box genes have been extensively studied in plants over decades (reviewed e.g. by Gramzow & Theissen 2010). These genes are best known for their role in the floral development, but they have also been shown to play a role in the development of other plant organs such as parts of flowers (Angenent & Colombo 1996), fruits (Liljegren *et al.* 1998, 2000) and roots (reviewed by Alvarez-Buylla *et al.* 2019).

Type I MADS box genes are less well known. These genes are similar in structure to the type II MADS box genes, but they lack the K domain. They can be further subdivided into Mα, Mβ, and Mγ subfamilies (Pařenicová *et al.* 2003). De Bodt *et al.* (2003a) made *in silico* analysis of Arabidopsis type I MADS box genes and discovered that most of them are expressed either at very low levels, or under highly specific conditions, and are thus under represented in EST collections. Function has been determined for a very few type I MADS box genes. *PHERES* is reported to have function in the early seed development of Arabidopsis (Kohler *et al.* 2003). The birth rate of type I MADS box genes in plants is higher than the birth rate of type II MADS box genes, but the latter are preferentially retained in genome (Nam *et al.* 2004). Masiero *et al*. (2011) reported important functions for the type I MADS box genes the during the reproductive development of Arabidopsis.
A short summary of some of the orthologuous of Arabidopsis A-, B-, C-, D- E-, FUL- and SOC-like genes in Petunia (*Petunia x hybrid*), Antirrhinum and Gerbera (*Gerbera hybrida*) is shown in table 1.

Table 1. Summary of selected orthologues of Arabidopsis A-, B-, C-, D- E-, FUL- and SOC-like genes in Petunia, Antirrhinum and Gerbera.

Type of Gene	Plant Species	Genes	References
Class A	Arabidopsis	*AP1, AP2*	Mandel et al. 1992, Jofuku et al. 1994
	Petunia	*PhAP2A, PhAP2B, PhAP2C, PFG*	Immink et al. 1999, Maes et al. 2001
	Antirrhinum	*SQUA, LIP1, LIP2*	Coen et al. 1990, Schwarz-Sommer et al. 1990
	Gerbera	N/A	
Class B	Arabidopsis	*AP3, PI*	Jack et al. 1992, Goto & Meyerowitz 1994
	Petunia	*PhGLO1, PhGLO2, PhDEF, PhTM6*	reviewed in Vandenbussche et al. 2004
	Antirrhinum	*DEF, GLO*	Schwartz-Sommer et al. 1990
	Gerbera	*GGLO, GDEF2, GDEF1*	Yu et al. 1999
Class C	Arabidopsis	*AG*	Bowman et al. 1991
	Petunia	*PMADS3, FBP6*	Tsuchimoto et al. 1993, Angenent et al. 1993
	Antirrhinum	*PLE, FAR*	Davies et al. 1999
	Gerbera	*GAGA1, GAGA2*	Yu et al. 1999
Class D	Arabidopsis	*STK, SHP1, SHP2*	Ma et al. 1991, Rounsley et al. 1995
	Petunia	*FBP7, FBP11*	Angenent et al. 1995, Colombo et al. 1995, Colombo et al. 1997
	Antirrhinum	N/A	
	Gerbera	N/A	
Class E	Arabidopsis	*SEP1, SEP2, SEP3, SEP4*	Ma et al. 1991, Mandel & Yanofsky 1998, Pelaz et al. 2000, Ditta et al. 2004
	Petunia	*FBP2, FBP5, PhAGL6, FPB9, FBP23, PMADS12*	Immink et al. 1999, Ferrario et al. 2004a
	Antirrhinum	*AmSEP3A, AmSEP3B, AmSEP3C, DEFH72, DEFH49, DEFH200*	Litt &Irish 2003
	Gerbera	*GRCD1,GRCD2,GRCD3,GRCD4,GRCD5*	Kotilainen et al. 2000, Uimari et al. 2004, Laitinen et al. 2005, II
FUL-like	Arabidopsis	*FUL, CAL*	Bowman et al. 1993, Mandel &Yanofsky 1995b, Kempin et al. 1995
	Petunia	*FBP29, FBP26*	Ferrario et al. 2004a, Immink et al. 1999
	Antirrhinum	*AmFUL, DEFH28*	Muller et al. 2001, Litt & Irish 2003
	Gerbera	*GSQUA1,GSQUA2,GSQUA3,GSQUA4,GSQUA5,GSQUA6*	Yu et al. 1999, Laitinen et al. 2005, I
SOC-like	Arabidopsis	*SOC1*	Borner et al. 2000
	Petunia	*UNS, FBP21, FBP28*	Immink et al. 2003, Vandenbussche et al. 2003a, Ferrario et al. 2004a
	Antirrhinum	*DEFH68*	Uniprot Q711P3
	Gerbera	*Gh-SOC1*	Laitinen et al. 2005, III

1.3 A function genes

Based on the ABC model, the criteria set for an A function gene are that they determine sepal and petal identity and repress C function in whorls one and two (Coen & Meyerowitz 1991). *Arabidopsis* has two A class genes *APETALA1* and *APETALA2* (*AP1*, *AP2*) (Irish & Sussex 1990, Mandel *et al*. 1992, Bowman *et al*. 1993, Kunst *et al*. 1989, Jofuku *et al*. 1994). Of these *AP1* is a MADS box gene (Jofuku *et al.* 1994). Interestingly, *AP1* was not considered an organ identity gene in the original version of the ABC model. Rather it was described in terms of its function in determining floral meristem identity (Coen & Meyerowitz 1991, Bowman *et al.* 1991, Meyerowitz *et al.* 1991). Only after the discovery of its expression pattern, which is restricted to whorls one and two by the repressive action of the C function gene *AGAMOUS* (*AG*), *AP1* was added to the model as an A function gene (Bowman *et al*. 1993, Gustafson-Brown *et al*. 1994). *AP1* has been shown to fulfill a dual function in specifying Arabidopsis sepal and petal identity and affecting inflorescence meristem development (Irish & Sussex 1990, Mandel & Yanofsky 1995b). *AP1* together with LFY controls flower meristem (FM) identity (Bowman *et al.* 1993). The expression pattern of *AP1* is in accordance with the ABC model by being expressed in whorls one and two (Mandel *et al*. 1992). *AP1* acts closely together and partially redundantly with other inflorescence architecture genes *CAULIFLOWER* (CAL) and

FRUITFULL (FUL) (Ferrándiz *et al.* 2000). Of these genes, only *AP2* meets both requirements for an A class gene. *AP2* is involved in sepal and petal identity specification and inhibits spreading of the C function to whorls one and two (Drews *et al.* 1991). *AP2* is the single ABC gene that is not a MADS box gene but belongs to AP2/ERF ethylene response family of plant transcription factors (Jofuku *et al.* 1994). *AP2* has been shown to have several other functions as well (reviewed by Litt 2007).

1.3.1 FUL-like genes

Functional homologues of Arabidopsis *AP1* have not been characterized in other plant species and that has raised questions whether A function itself is limited to Arabidopsis or to Brassicaceae (reviewed by Litt 2007). Paralogous genes to *AP1* have risen as a result of gene duplication in Arabidopsis. These genes are similar in sequence and partially redundant in action. One such gene is *CAULIFLOWER* (*CAL*), which specifies FM identity but has no function in sepal or petal identity determination (Bowman *et al*. 1993, Kempin *et al*. 1995). However, in the absence of *AP1* function, *CAL* might be involved in petal identity determination (Castillejo *et al*. 2005).

Arabidopsis *FRUITFULL (FUL)* that shares high sequence identity with *AP1*, is reported to have an early function in controlling flowering time, meristem identity and cauline leaf morphology. Later during development, *FUL* is involved in carpel and fruit development affecting valve, replum and style morphology (Mandel & Yanofsky 1995a , Gu *et al*. 1998, Ferrándiz *et al.* 2000, Liljegren e*t al.* 2000, 2004, Ferrándiz 2002). *FUL*-like genes have been identified from several species besides Arabidopsis, eg. Antirrhinum, orchid *Dendrobium thyrsiflorum*, birch *Betula pendula*, bilberry *Vaccinium myrtillus* and poppies *Papaver somniferum* and *Eschscholzia californica* (Muller *et al*. 2001, Skipper *et al.* 2005, Elo *et al.* 2001, Jaakola *et al.* 2010, Pabón-Mora *et al*. 2012). Some of these genes have earlier been misinterpreted as *AP1*-like based simply on sequence similarity.

FUL-like genes are expressed in wide range of floral organs and vegetative parts. They are commonly expressed in carpels (Immink *et al*. 1999, 2003, Wu *et al*. 2000, Sung *et al.* 2001, Jang *et al*. 2002, Busi *et al*. 2003, Hileman *et al.* 2006), meristems (Mandel & Yanofsky 1995a, Gu *et al*. 1998) and vegetative tissue including bracts (Sung *et al*. 1999, Immink *et al*. 2003, Sreekantan *et al.* 2004). Expression has also been detected in inflorescence and floral meristem, stamens and perianth organs (Almeida *et al*. 2015). For some species expression has been visible in all floral whorls (eg. Pabón-Mora *et al.* 2012).

Sequences of *AP1*- and *FUL*-like genes are very similar and based solely on sequence comparisons, many genes have been misclassified as *AP1*-like when they are closer related to the *AP1* paralogs

CAL and *FUL*. The status of *AP1*- and *FUL*-like genes was further specified by Litt & Irish (2003) and Vandenbussche *et al.* (2003b) as they described the conserved motifs at the C terminus of these proteins. They were named euAP1 and paleoAP1 motifs (Vandenbussche *et al.* 2003b), the latter is called FUL-like motif by Litt & Irish (2003). These motifs greatly facilitate identification of the related proteins as detailed phylogenetic analysis is not an absolute requirement. The paleoAP1/FUL-like protein sequences have a hydrophobic box (L/MPPWML), which is not found in euAP1-like sequences. EuAP1-like sequences have two conserved motifs, a transcription activation domain RNNaLaLT/NLa for euAP1 motif, and a farnesylation signal (CFAT/A), which terminates the protein. The farnelysation motif generally directs proteins to a membrane (Hancock *et al.* 1991). In case of transcription factors this can be part of posttranscriptional regulation, or necessary for protein complex formation (Yalovsky *et al.* 2000b). AP1 has been shown to be farnesylated *in planta*, but membrane localization has not been detected (Yalovsky *et al.* 2000b). Yalovsky *et al.* (2000a) speculate that farnesylation may facilitate interaction of AP1 with the other developmentally regulated transcription factors. Not all euAP1-like proteins possess the farnesylation signal (eg. Berbel *et al.* 2001), and thus it may not be an essential part of the protein function (Litt 2007). Functional domains have been found only in the euAP1 proteins, and thus it has been speculated that they provide functional capabilities, which might not exist in FUL-like proteins. Later studies, however, suggest that this is probably not true, as the novel C terminal motifs do not assign novel roles for these proteins. Studies on poppies raise the possibility that regulation rather than sequence is the key player in role determination for euAP1- and FUL-like proteins (Pabón-Mora *et al.* 2012). In their studies Litt & Irish (2003) and Vandenbussche *et al.* (2003b) described protein motifs for several classes of MADS domain proteins besides euAP1- and FUL-like proteins.

1.3.2 The existence of A function outside Arabidopsis

SQUAMOSA (*SQUA*) of Antirrhinum is an *AP1* orthologue. *SQUA* acts in determining floral meristem identity, but no role in sepal and petal identity specification has been reported. Neither is *SQUA* able to repress *PLENA*, the Antirrhinum C class gene, in whorls one and two (Huijser *et al.* 1992). Two *FUL* orthologues, *AmFUL* and *DEFH28*, have been identified in Antirrhinum (Litt & Irish 2003, Muller *et al*. 2001). *AmFUL* has not been functionally characterized (Litt & Irish 2003), but the expression was analysed by Preston & Hileman (2010). *AmFUL* expression was detected in sepals, petals and gynoecia, in more detail the expression was seen in the apical inflorescence meristem and lateral floral meristems (Preston & Hileman 2010). Based on its expression pattern and phenotypic changes in flowers and fruits (Muller *et al*. 2001), *DEFH28* is speculated to play a similar role in Antirrhinum as *FUL* does in Arabidopsis (Litt 2007). Antirrhinum *AP2*-like genes *LIP1* and *LIP2* are essential for sepal and petal development but do not repress C class gene function in whorls one and two (Keck *et al.* 2003).

Petunia has three *AP2*-like genes, *PhAP2A*, *PhAP2B* and *PhAP2C*. Of these genes *PhAP2A* shares high sequence similarity and similar expression pattern with Arabidopsis *AP2* suggesting orthology. Functional orthology has been shown by complementation studies (Maes *et al*. 2001). *PhAP2B* and *PhAP2C* belong to a different subgroup of the *AP2* family and their expression diverges from that of *PhAP2A* (Maes *et al*. 2001). The suspected Petunia A function gene *blind* (Vallade *et al.* 1987, Maes *et al*. 2001) was in fact found to encode a micro-RNA regulating C function (Cartolano *et al.* 2007). No gene orthologous to the Arabidopsis *AP1* has been discovered in Petunia, and the results suggest that the A function in Petunia is organized differently compared to Arabidopsis (Rijpkema *et al.* 2006).

Despite of the attempts to prove a similar function for related genes in other plant species, success has been scarce. In many cases meristem maintenance function has been shown to apply, but involvement in specifying sepal and petal identity has been negative. For example, the Antirrhinum *SQUAMOSA* (*SQUA)* plays a role in inflorescence meristem development but does not affect flower organ identity (Huijser *et al.* 1992). The same function has been shown for the Anthirrinum gene *DEFH28*, which is not involved in sepal and petal identity determination (Muller *et al*. 2001). Several plant species (e.g. *Eucalyptus* spp., *Betula pendula*, *Solanum tuberosum*, *Oryza sativa*, *Dendrathema grandiflorum*, *Salix discolor*) have been shown to have genes closely related to *AP1* but none have similar function in specifying sepal and petal identity (Kyozuka *et al*. 1997, Elo *et al*. 2001, Hart & Hannapel 2002, Fornara *et al*. 2004, Shchennikova *et al.* 2004, Fernando & Zhang 2006). Some of these genes have been falsely interpreted as *AP1*-like based on sequence similarity without phylogenetic data. After closer analysis the genes were identified as *FUL*-like. These genes are not likely to have a function in organ identity determination, but rather have a role in specifying floral meristem identity and in affecting flowering time. The pea (*Pisum sativum*) MADS box gene *PEAM4* seems to be the closest to *AP1* in function and has been suggested to be a functional homologue of *AP1* (Berbel *et al*. 2001) based on similar expression pattern and floral phenotype. The *EuAP1*-like genes have been studied more than the *paleoAP1/FUL*-like genes whose functions are just emerging (e.g. Pabón-Mora *et al*. 2012).

1.4 B class genes

The classical B class genes of Arabidopsis are *PISTILLATA* (*PI*) (Goto and Meyerowitz 1994) and *APETALA3* (*AP3*) (Jack *et al*. 1992). They are both active in whorls two and three, specifying petal and stamen identity. The origin of the paralogous B gene pair is thought to be a duplication event that occurred before the origin of the angiosperms (Kramer *et al.* 1998, Purugganan 1997, Theissen *et al.* 2000, Kim *et al.* 2004). *AP3* and *PI* form heterodimers (Riechmann *et al*. 1996a, Riechmann *et*

al. 1996b), which is important for their regulation as dimerization enhances and maintains their expression (Honma & Goto 2000). Based on the C-terminal motifs, two clades of AP3-like proteins can be distinguished. The paleoAP3 motif is found from basal eudicots, magnoliids, monocots, and basal angiosperms. Respectively, the euAP3 motif is only present in core eudicots (Kramer *et al.* 1998). Some core eudicot species harbour both types of genes, called *euAP3*-like and *TOMATO MADS BOX GENE6* -like (*TM6*-like), respectively (Kramer & Irish 2000). Arabidopsis and Antirrhinum have lost the paleoAP3 gene. The EuAP3 motif is most likely a result from a simple frameshift mutation in one of the duplicated *paleoAP3*-genes (Vandenbussche *et al.* 2003b). The different C terminal motifs could encode different functions (Lamb & Irish 2003), but also contradictory results have been published (Whipple *et al.* 2004, Pabón-Mora *et al.* 2012). The overall functional significance of these motifs is not yet clear.

The Antirrhinum B class gene orthologues are *DEFICIENS* (*DEF*) and *GLOBOSA* (*GLO*) (Sommer *et al.* 1990, Tröbner *et al.* 1992, Schwarz-Sommer *et al.* 1992). They also form a heterodimer which leads to the stabilization of protein complex, and an autoregulatory circuit to maintain their expression. They are expressed in the second and third whorl of the flower.

In contrast to Arabidopsis and Antirrhinum, Petunia has two *PI* linage genes, *PhGLO1* (aka *FBP1*) and *PhGLO2* (aka *PMADS2*/*FBP3*). The expression patterns are comparable to the orthologous genes in Arabidopsis and Antirrhinum being confined to whorls two and three (Angenent *et al*. 1992, van der Krol *et al.* 1993, Vandenbussche *et al*. 2004). Despite some differences in function, *PhGLO1* and *PhGLO2* act largely redundantly in petal and stamen formation (Vandenbussche *et al.* 2004).
Petunia differs from Arabidopsis and Antirrhinum also in the case of *AP3*-like genes. Both Arabidopsis and Antirrhinum *AP3* and *DEF* belong to the modern clade of *AP3* lineage and harbour the euAP3 motif. Petunia however has both versions of the gene, *PhDEF* (aka *GREEN PETAL* (*GP*)/*PMADS1*) and *PhTM6* (Angenent *et al*. 1992, van der Krol *et al.* 1993, Kramer & Irish 2000, Vandenbussche *et al*. 2004). *PhDEF* has been shown to be essential for petal identity, but it might act redundantly with other factors in stamen development (Vandenbussche *et al*. 2004).

While the other Petunia B class gene orthologs are very similar to the corresponding genes in Arabidopsis and Antirrhinum, *PhTM6* differs drastically from them (Vandenbussche *et al.* 2004). The highest expression of *PhTM6* is detected in carpels and stamens, while expression in sepals and petals is clearly lower. Later in development, the expression remains high in carpel and declines in stamens as they reach maturity. The high expression level of both *PhDEF* and *PhTM6* in anthers suggests redundant function in stamen formation (Vandenbussche *et al.* 2004). *PhTM6* is regulated like a C class gene rather than a B class gene (Rijpkema *et al*. 2006). One or both of Petunia C

function genes are required for the maintenance of *PhTM6* expression in Petunia (Heijmans *et al.* 2012).

The ABC model explains reasonably well floral structures in core eudicots but extending the model to the monocots and many basal angiosperms, brings out the tremendous diversity in floral structures and organization (Friis *et al.* 2000). Several alternative models have been suggested based on floral morphology, and/or patterns of gene expression (van Tunen *et al.* 1993, Bowman 1997, Kanno *et al*. 2003, Kramer *et al.* 2003, Buzgo *et al*. 2004, Kim *et al.* 2005). The extended B function model (the tulip/lily model) explains the existence of two petaloid whorls found in *Lilium* by the expression of B genes in whorl one. The sepals of the first whorl are converted to the petals (van Tunen *et al*. 1993). Most tulips also possess two whorls of perianth organs (tepals) similar to *Lilium,* which express the B function homologs (Kanno *et al.* 2003). The same phenomenon has been described in more general terms as the shifting boundary (Bowman 1997) and sliding boundary model (Kramer *et al.* 2003). In this model the B function moves across the floral meristem and is active in the whorl one. An inward shift in the A function in turn might result in the petaloid stamens and/or staminodes. This is not allowed by the classical ABC model since the A and C functions are assumed to be antagonistic (Coen & Meyerowitz 1991). The gradual transition from the outer to the inner floral organs characterizes two of the three basal lineages of angiosperms, *Amborella* and Austrobaileyales and might to be the ancestral condition (Ronse de Craene *et al.* 2003, Soltis *et al.* 2005). Buzgo *et al*. (2004) suggested that this gradual transition in the organ morphology across the floral meristem is caused by a gradient in the floral regulator expression levels. Weak expression at the outer borders of a given gene's range of activity overlaps with the neighboring gene's expression and thus centers of strong expression are separated from each other. This way adjacent organs are affected by each other, and produce mosaic-like organs, rather than clear-cut distinct organs. Expression of the B function homologous genes in *Amborella* supports this model (Kim *et al.* 2005).
The classical ABC model and B function modified versions are shown in figure 2.

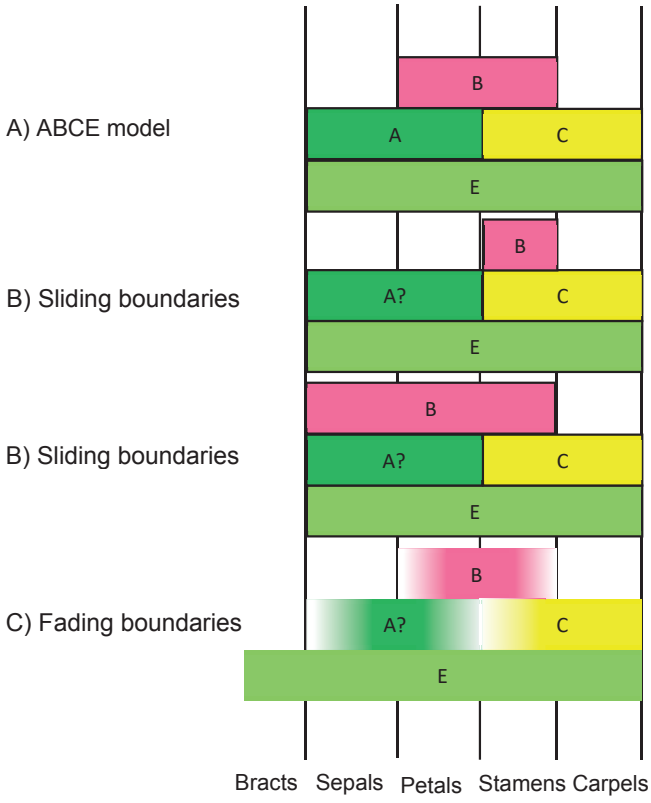

Figure 2. The Classical ABCE model and the extended B models for floral organ identity determination (modified from Soltis *et al.* 2007). The classical model (Coen & Meyerowitz 1991, Colombo *et al*. 1995, Pelaz *et al.* 2000). The sliding/shifting boundary model (van Tunen *et al.* 1993, Bowman 1997, Kramer *et al*. 2003). The fading boundary model (Buzgo *et al.* 2004).

1.5 C class genes

The C function in Arabidopsis is encoded by *AGAMOUS* (*AG*) (Yanofsky *et al.* 1990). Besides conferring stamen and carpel identity, it maintains flower meristem identity. C function is antagonistic to A function and in the absence of A function, it spreads to whorls one and two. When *AG* expression is downregulated, stamens convert into petals, and carpels are replaced by indeterminate perianth whorls (Yanofsky *et al.* 1990). *AG* affects flower development by controlling other regulatory genes,

which are involved in many different developmental processes. The *AG* action suppresses the leaf development program in the emerging floral primordia (Ó'Maoiléidigh *et al*. 2013).

In Antirrhinum, C function is governed by two genes, *PLENA* (*PLE*) and *FARINELLI* (*FAR*). These proteins are highly similar, most likely as the result of a duplication event. Despite their sequence similarity they play distinct and separable roles. Downregulation of *PLE* results in a phenotype, which closely resembles that of the downregulated *AG,* although the protein sequence of FAR is closer to AG than PLE is (Davies *et al.* 1999). In fact, PLE appears to be the ortholog of SHATTERPROOF protein (D function, see later) (Kramer *et al.* 2004). Downregulation lines for *PLE* and *FAR* genes show that while *FAR* in the stamen development can be substituted by the redundant action of *PLE*, the opposite does not work. The reason is the complex regulatory relationship between these two C function genes (reviewed in Davies *et al.* 2006). The case of *PLE* and *FAR* has been interpreted to represent a situation where, after gene duplication, a different copy of the paralogous gene pair has retained the ancestral developmental function (Causier *et al.* 2005).

In Petunia, C function is encoded by two genes, Petunia *MADS3* (*pMADS3*) and *FLORAL BINDING PROTEIN 6* (*FBP6*) (Angenent *et al*. 1993, Tsuchimoto *et al.* 1993). Both these sequences are homologous to *AG* of Arabidopsis. At very early stages of development, the transcripts become visible in cells, which later give rise to stamen and carpel primordia. Only at later stages of development, the genes are differentially expressed. *FBP6* is most highly expressed in the stigma and transmitting tissue of the style, while *pMADS3* expression is mostly detected in ovules, vascular tissue and nectaries (Kater *et al.* 1998). Based on several studies (Tsuchimoto *et al*. 1993, Kater *et al.* 1998, Kapoor *et al.* 2002), *pMADS3* is considered the Petunia ortholog of *AG* and is thus required for stamen and carpel development. Recent studies showed *pMADS3* and *FBP6* to be redundant in floral organ identity and floral determinacy in contrast to subfunctionalization visible both in Arabidopsis and Antirrhinum (Heijmans *et al*. 2012).

1.6 D class genes

Originally the D function genes were described in Petunia by Angenent *et al.* and Colombo *et al.* in 1995. The Petunia D function genes are called *FLORAL BINDING PROTEIN7* (*FBP7*) and *FLORAL BINDING PROTEIN11* (*FBP11*). At the protein level *FBP7* and *FBP11* are 90% similar, and they share high sequence similarity with *SEEDSTICK* (*STK*) of Arabidopsis (Angenent *et al.* 1995, Pinyopich *et al.* 2003). Both these genes are expressed in the centre of the gynoecium before the ovule primordia emerge. At later stages, the expression is restricted to ovules and the expression levels of both genes peak right after pollination, then declining in the developing seeds (Colombo *et al.* 1995, 1997). However, the knock-down mutants for both genes did not show significant alterations

in ovule development, and thus the authors concluded that Petunia ovule development is not entirely specified by *FBP7* and *FBP11,* but rather affected by all Petunia *AG*-like genes (Heijmans *et al*. 2012). Petunia D function genes together with Petunia C function genes were also discovered to participate in floral termination, as suggested also by earlier studies (Ferrario *et al.* 2006, Heijmans *et al.* 2012).

In Arabidopsis, the *AG* subfamily includes the D function genes *SHATTERPROOF1/2* (*SHP1/2*) and *SEEDSTICK* (*STK*) (Favaro *et al.* 2003, Pinyopich *et al.* 2003), which define ovule identity. Moreover, these genes also have functions in fruit development and seed dispersal (Liljegren *et al.* 2000, Pinyopich *et al*. 2003). In Antirrhinum some results suggest that PLE (a C function protein) to be an ortholog of SHP proteins (Kramer *et al.* 2004).

1.7 E class genes

E function in Arabidopsis is encoded by four *SEPALLATA* genes (*SEP1/2/3/4*) (Ma *et al*. 1991, Huang *et al.* 1995, Pelaz *et al.* 2000, Ditta *et al.* 2004), which were previously known as *AGL2, AGL4, AGL9* and *AGL3*. These genes were found in the early 1990s, but they were characterized considerably later. A triple knock-down mutant (*sep1/sep2/sep3*) was required to reveal their function due to their redundant nature (Pelaz *et al*. 2000). The triple mutant showed that B and C organ identity functions require SEP1, SEP2 and SEP3 in order to form petals, stamens and carpels. Further studies and a quadruple mutant (*sep1/sep2/sep3/sep4*) indicated the involvement of all SEP genes in specifying sepal identity. *SEP4* also participates in petal, stamen, carpel and meristem identity specification (Ditta *et al*. 2004). *SEP3* in the best characterized among the Arabidopsis *SEP* genes (Ditta *et al*. 2004, Honma & Goto 2001, Pelaz *et al*. 2001a, Favaro *et al*. 2003, Castillejo *et al*. 2005) and new evidence questions the presumed high redundancy among the *SEP* genes. *SEP3* has a more prominent role in Arabidopsis floral development than other *SEP*s by influencing floral meristem identity, floral organ identity, and ovule development (Favaro *et al*. 2003, Pelaz *et al*. 2001a, Ditta *et al.* 2004, Honma & Goto 2001, Zahn *et al*. 2005).

Six *SEP clade* genes (*FBP2, FBP4, FBP23, FBP3, FBP9* and *pMADS12*) have been identified in Petunia (Angenent *et al.* 1992, Ferrario *et al.* 2003, Immink *et al.* 2003, Vandenbussche *et al.* 2003a). *FBP2*, *FBP5*, *FPB23* and *pMADS12* are flower specific, while *FBP4* and *FBP9* are expressed also outside the floral organs (Ferrario *et al*. 2003). At least two genes, *FBP2*, and *FBP5* contribute to E function in Petunia. When these genes are co-suppressed, three outermost floral whorls lose their identity (Angenent *et* al. 1994, Ferrario *et al.* 2003). Double knockdown mutant plant phenotype, however, suggests the presence of at least one more partially redundant E function gene (Vandenbussche *et al*. 2003a). A recent study by Morel *et al.* (2019) further investigated Petunia

SEP/AGL6 and *AP1/SQUA* genes. According to this study, Petunia *SEP1/SEP2/SEP3* orthologs and *ALG6* together encode the classical *SEP*-function of floral organ identity and floral determination function present also in Arabidopsis (Angenent *et al*. 1994, Pelaz *et al*. 2000, Honma & Goto 2001). The major role in determining these functions is played by the *SEP3* ortholog *FBP2*. The floral meristem identity in Petunia is controlled by *FBP9* and *FBP23* together with *FBP4*, but these genes do not contribute to floral organ identity much. Interestingly, *FBP9* and *FBP23* do not have clear orthologs in Arabidopsis. The four Petunia *AP1/SQUA*-like genes characterized in this study are required for floral meristem identity and they act redundantly. These genes also suppress the B function in the first floral whorl of Petunia (Morel *et al.* 2019).

1.8 *SOC1*-like genes

Arabidopsis *SOC1* is the best characterized member of this class of the MADS box genes. As flowering time integrator, this gene combines several pathways (photoperiodic, vernalization, energy, and gibberellin) leading to flowering (reviewed by Lee & Lee 2010). *SOC1*-like genes have been shown to function as floral integrators also in other plant species (Immink *et al.* 1999, Borner *et al*. 2000, Ferrario *et al*. 2004, Lee & Lee 2010, III). A role in floral patterning and floral meristem identity determinacy has also been assigned to *SOC1* (Liu *et al*. 2007, 2009, Melzer *et al*. 2008). In addition to the floral associated function, *SOC1*-like genes are involved in vernalization and senescence processes (Zhang & Forde 1998, Samach *et al*. 2000, Ratcliffe *et al*. 2001, Trevaskis *et al*. 2003, Fang & Fernandez 2002, Vrebalov *et al.* 2002, Ratcliffe *et al.* 2003). Other, less characterized paralogous genes of Arabidopsis include *AGL14, AGL19, AGL42, AGL71*, and *AGL72* (Becker & Theissen 2003).

SOC1-like genes are a group of MADS box genes, which has been suggested by some authors to be called the F function genes in reference to the classical ABC model (Coen and Meyerowitz, 1991; *Nam et al*., 2003; Nam *et al*., 2004). Their expression pattern is preferentially vegetative in both gymnosperm and angiosperm species (Tandre *et al*.1995; Walden *et al.* 1998; Winter *et al*. 1999; Watson & Brill 2004), but expression is seen also in reproductive organs (Heuer *et al.* 2001; Münster *et al*. 2002). *SOC1*-like genes are associated with vascular development and wood formation in Eucalyptus, Arabidopsis and Populus (*Populus tremuloides*) (Decroocq *et al.* 1999; Alverez-Buylla *et al.* 2000a, Cseke *et al.* 2003). Overexpression of *SOC1*-like genes typically causes early flowering (Tadege *et al*. 2003; Ferrario *et al.* 2004; Smykal *et al.* 2007).

The best characterized *SOC1*-like gene of Petunia is called *UNSHAVEN* (*UNS*) (Ferrario *et al.* 2004). Other Petunia genes in the *SOC1* class include *FBP21*, *FBP22*, and *FBP28.* All of them are expressed mainly in the vegetative parts, with some variation reported (Immink *et al.* 2003). Both

UNS and *FBP28* were differentially expressed but failed to produce phenotypes. However, ectopic expression of *UNS* led to early flowering and vegetative characteristics within the flower (Ferrario *et al*. 2004, Vandenbussche *et al*. 2003a).

1.9 MADS box gene duplication and redundancy

Large gene families are formed by gene duplication events. Duplications can occur at the level of whole genomes, segmental parts of genomes, or single genes (reviewed in Panchy *et al.* 2016). After duplication, both gene copies can be retained in the genome, or one of the paralogous genes acquires deleterious mutations and becomes a pseudogene (nonfunctionalization). This is the most common fate of a duplicated gene. After a gene duplication, both copies can be retained in the genome, if they are recruited for different functions (subfunctionalization). Duplicate genes partition the original function and both copies retain a subset of the function.Duplicated genes can also exist in a partially redundant state over a long evolutionary time. This is visible in the Arabidopsis C and D function genes. The Arabidopsis C function genes *AG* and the D function *STK, SHP1* and *SHP2* are all involved in the determination of the carpel identity, which is their shared ancestral role (Pinyopich *et al.* 2003). Lineage specific duplication of the maize *AGAMOUS* -like gene (responsible for formation of stamens and carpels) is also an example of subfunctionalization (Mena *et al.* 1996). The ancestral function of *AG* is split and *ZMM1* encodes stamen function and *ZAG1* encodes carpel function (Mena *et al.* 1996). In some circumstances, purifying selection constrains the divergence between redundant genes (Moore *et al.* 2005). Arabidopsis SEP genes are duplicated genes with partially redundant function (Pelaz *et al*. 2000, Ditta *et al*. 2004). The single gene knock down mutants only show very mild phenotypes, or no phenotype is observed. Only after triple and quadruple mutants, a dramatic phenotype appears. In Gerbera paralogous genes *SEP*-like *GRCD1* and *GRCD2* display similar subfunctionalization event. *GRCD1* is active in stamen formation, whereas the closely related *GRCD2* has a specific function in carpel development (Kotilainen *et al*. 2000, Uimari *et al*. 2004).

The likelihood of a gene being retained in the genome after the whole genome duplication varies, and the gene balance theory was described by Edger & Pires (2009). Genes coding proteins that are highly connected in networks, e.g. transcription factors, signalling and posttransciptional protein modification proteins, are preferentially retained as duplicates. The loss of such genes might be deleterious by perturbing the balance of e.g. protein complex formation and the subsequent regulation of target genes. Due to dosage effect, after small-scale duplications, also the genes coding for highly connected proteins are preferentially lost. If only a single gene, or a small segment is duplicated, it disturbs the balance of the biological processes downstream (Edger & Pires 2009).

The duplicated genes can reach an evolutionary stable equilibrium of partial redundancy, which might delay the subsequent complete subfunctionalization (Wagner 1999). Redundancy was considered evolutionarily unstable due to divergence of duplicated genes that increases robustness and evolvability of an organism (reviewed in Kafri *et al.* 2009). Yet, not all redundant genes are unstable, but can be maintained in genome for 80-100 million years (e.g. Tischler *et al*. 2006, Dean *et al*. 2008, DeLuna *et al*. 2008). The long evolutionary co-existence of these duplicate genes must reflect their importance to organism fitness (Kafri *et al.* 2009). Three models for stable co-existence of duplicate genes were first described by Nowak *et al.* (1997). Selective advantage created by gene duplication is another possibility for their retention in genome (Kafri *et al.* 2009). When both copies are retained in the genome, even if partially redundant, the advantage for the plant is increased mutational robustness that leads to evolutionary innovation e.g. flowering plants (Wagner 2008). Whole genome duplications can facilitate speciation by creating a reproductive isolation (Scannell *et al.* 2006, Edger & Pires 2009). Another outcome of a whole genome duplication is raw material for mutation and selection, which generates new features and increases complexity (Freeling & Thomas 2006, van de Peer *et al*. 2009).

If the duplicated gene acquires a new function, the event is called neofuctionalization. Neofunctionalization is a rare event and requires mutations either in the regulatory, or the coding region of the gene. New functions are likely to arise at the late phase of evolution when the paralogous sequence has acquired a higher level of divergence (Kondrashov *et al*. 2002). Functional diversification can occur at the regulatory regions of genes, which affects the temporal, spatial and/or quantitative expression. Alterations within coding region change a protein function (Blanc & Wolfe 2004). The Arabidopsis *AP1/FUL*-like genes illustrate partial redundancy and neofunctionalization. *AP1*, *CAL* and *FUL* act redundantly to specify floral meristem identity in transition from vegetative stage to inflorescence development (Kempin *et al.* 1995, Mandel & Yanofsky 1995a, b, Ferrandiz *et al.* 2000). Besides that, AP1 and FUL have also non-redundant functions (Gu *et al.* 1998, Liljegren *et al.* 2004). Neofunctionalization is nicely illustrated by chinese lantern (*Physalis floridana*) (He & Saedler 2005), where *MPF2* (a MADS box gene) heterotopic expression in flower contributed to the development of new morphological trait known as inflated calyx syndrome (He *et al*. 2004).

Duplication events play a major role in driving evolutionary diversification of plant genomes. Duplicated loci compose a prominent portion of plant genomes. Both segmental genome duplications and whole genome duplication events have contributed to apparent redundancy. According to Jiao *et al*. (2011) a whole genome duplication (WGD) took place in the common ancestor of all seed plants 320 million years ago. Second WGD happened 190 million years ago at the origin of angiosperms. After the diversification of monocot and dicots approximately 150 million years ago, each lineage had several WGDs (Jiao *et al*. 2011). Bowers *et al.* (2003) have suggested

that Arabidopsis has undergone three whole genome duplication events. In these events genes with regulatory function (e.g. transcription factors) have been retained more often than genes originating from single gene duplication as the gene balance theory suggests (Maere *et al.* 2005, Edger & Pires 2009). Interestingly also, Arabidopsis seems to retain duplicated genes much longer than animals (Lynch & Conery 2001). The basis for this phenomenon is not known, and its applicability to other species is not clear.

Whole genome duplications are considered as regular events in plant evolution (Causier *et al.* 2005, Zahn *et al.* 2005) and they have given rise to the large number to type II MADS box genes (MIKC MADS box genes) controlling flower development. Many type II genes were preferentially retained after duplication events, gene sequences diverged and resulted in sub- and neofunctionalization (Gramzow & Theissen 2013, 2015, Theissen & Gramzow 2016). The exceptionally high number of type II genes is specific to plants (Nam *et al*. 2003). Type I MADS box genes are suggested by Nam *et al.* (2004) to be the outcome of the smaller, local duplication events. The birth rate of type I genes is higher than that of type II genes, but the latter are preferentially retained in genome (Nam *et al.* 2004).

Gramzow *et al.* (2010) have suggested a theory for the early evolution of MADS box genes. Based on the sequence similarity of the DNA binding domain of MADS transcription factors with A subunit of topoisomerase IIA, they suggest that TOPOIIA-A duplicated in the common ancestor of all eukaryotes. One copy retained the original function necessary e.g. for DNA replication, and the other copy acquired the DNA binding capacity and became the progenitor of MADS transcription factor (Gramzow *et al.* 2010). MADS box genes were further divided into two types (I and II) based on their structure (Alvarez-Buylla *et al.* 2000). Orthologous genes (SFR-like and MEF2-like) are found in animals. This refers to a much later duplication in the common ancestor of plants and animals.

Generally, MADS box genes in different plant species show a rich history of different fates after duplication events. In many cases, it is a combination of gene loss, neofunctionalization, subfunctionalization, or the mixture of them all (reviewed in Airoldi & Davies 2012).

1.10 Higher order complexes of MADS domain proteins

MADS domain proteins act in a combinatorial manner to specify floral organ identity. The first evidence for multimeric complex formation of these proteins was shown in Antirrhinum by Egea-Cortines *et al.* (1999) and in Arabidopsis by Honma and Goto (2001).

MADS domain transcription factors bind DNA sequence specifically, with the core consensus 5'-$CC(A/T)_6GG$-3'. This consensus sequence is termed CArG box (Passmore *et al.* 1989, Pollock & Treisman 1990, Wynne & Treisman 1992). The core CArG box is 10 nucleotides long with some of the surrounding nucleotides conserved, and the whole consensus sequence consists of 16 nucleotides. MADS domain proteins bind to the CArG-box as either homo- or heterodimers (reviewed in Kaufmann *et al*. 2005). The binding affinity of the MADS domain proteins to the consensus sequence varies, and the target sequence can vary slightly (de Folter & Angenent 2006, Yan *et al*. 2016). A recent study by Aerts *et al.* (2018) shows that the definition of the CArG box is more flexible than described earlier, and it may differ from the classical consensus sequence. Different sets of CArG boxes were enriched for different MADS domain proteins and the boxes containing an extension NAA were clearly preferred. Motifs of other transcription factors were found in the vicinity of the CArG boxes, which highlights the importance of interactions of MADS domain proteins with other transcription factors in target gene regulation (Aerts *et al*. 2018). Several factors, like chromatin condensation, presence of co-factors and competing proteins, influence whether a certain CArG-box is bound by a MADS domain protein (Wray *et al.* 2003).

Plant MADS domain proteins have been shown to form higher order complexes, which are larger than dimers (Egea-Cortines *et al.* 1999, Honma & Goto 2001). Based on these results, a quartet model was proposed (Theissen 2001, Theissen & Saedler 2001). This model proposes that the tetramers consisting of two MADS domain protein dimers, determine floral organ identity. The tetrameric complexes, which consist of two dimers of MADS domain proteins, bind to two CArG boxes located 10 to 200 bp apart (Kaufmann *et al.* 2009). DNA sequence is looped between the bound dimers (Theissen 2001, Theissen & Saedler 2001). According to the ABC model (Coen & Meyerowitz 1991) Arabidopsis sepals are specified by two dimers of AP1 and SEP. AP1 and SEP dimer together with B function dimer PI and AP3 gives rise to petals in whorl two. Stamens are formed by the action of SEP and AG dimer with the B function dimer. Two pairs of SEP and AG dimers determine the formation of carpel in the center of the flower (Theissen 2001, Theissen & Saedler 2001, Krizek & Fletcher 2005). The central role of SEP proteins is illustrated by its presence in all regulatory complexes described above, as well as shown or predicted presence in complexes specifying several other organs or tissues (Favaro *et al.* 2003, Kaufmann *et al*. 2005, Castillejo *et al.* 2005).

The floral quartet model has been largely accepted, but experimental evidence supporting it has been limited. One of the few studies addressing this dilemma was published by Melzer *et al.* in 2009. In this elegant experiment they showed that *in vitro* homotetramer of Arabidopsis SEP3 binds to CArG boxes. They also showed that truncated version of SEP3 lacking K3 and C terminal domain was capable of dimerization, but not able to form tetramer, and thus bind DNA (Melzer *et al.* 2009).

They suggest that spacing and phasing of GArG boxes is of relevant importance for the molecular mechanism, which controls floral homeotic proteins and their target gene specificity. This study also suggests that quartet formation does not require the interaction of different floral homeotic proteins, rather the intrinsic capacity of SEP3 to tetramerize is sufficient.

Yeast based interaction studies on MADS domain protein-protein interaction facilitating higher order complex formation have been performed on several plant species like Gerbera, Petunia, Chrysanthemum, rice, perennial ryegrass, tomato etc. (Kotilainen *et al*. 2000, Uimari *et al*. 2004, Immink *et al.* 2003, Shchennikova *et al.* 2004, Fornara *et al.* 2004, Ciannamea *et al.* 2006, Leseberg *et al.* 2008). For Arabidopsis an extensive screen of all the MADS domain protein has been carried out (de Folter *et al.* 2005).

The floral quartet model interactions of the MADS domain proteins has been confirmed. The MADS domain protein complexes are flexible in nature depending on the concentration of the proteins and the DNA sequence. The chromatin remodelling factors and other transcription factors have been identified as interaction partners. MADS domain proteins participate in large protein complexes and act in a combinatorial manner with other non-MADS transcriptional regulators (Smaczniak *et al.* 2012).

1.11 *Gerbera hybrida*, the model plant of sunflower family

Arabidopsis has been the traditional model species for developmental studies. However, several authors are pointing out the challenges of extrapolating results from *Arabidopsis* directly to the other species including the generality of the ABC model (Litt & Irish 2003, Davies *et al*. 2006, Litt 2007). Moreover, the diversity of floral traits, or even gene copies that are not present in Arabidopsis, necessitates gene functional studies in other model species.

Gerbera hybrida is a representative of the Asteraceae plant family consisting of over 23 000 species worldwide (Jeffrey 2007). Typical to the sunflower family Gerbera differs from the other common plant model species by exhibiting a composite inflorescence (head or capitulum). The complex floral system bears individual flowers and, resembling a large single flower, attracts pollinators. The genetic network regulating the development of such a complex structure is likely to differ from the systems controlling the development of simpler floral structures of Arabidopsis, Antirrhinum and Petunia.

The floral head (layman's flower) consists of several hundreds of the specialized flower types. In the inflorescence three different flower types can be differentiated (Harris 1995). The marginal ray

flowers are largest in size and form the pseudanthium mimicking the perianth of a single flower. These flowers are strongly zygomorphic consisting of three fused petals which form a ligule, two rudimentary petals and a carpel. The flowers in the center, the disc flowers are more radially symmetrical, and bear a less distinctive corolla. Disc flowers are the only perfect flowers of Gerbera capitulum, containing also stamens in addition to the carpel. The third flower type resides between the ray and disc flowers and is called the trans flowers. The trans flowers share features with both the flower types surrounding it. They are female flowers like the ray flowers, but the corolla is shorter like in the disc flowers. The structure of Gerbera inflorescence is shown in figure 3.

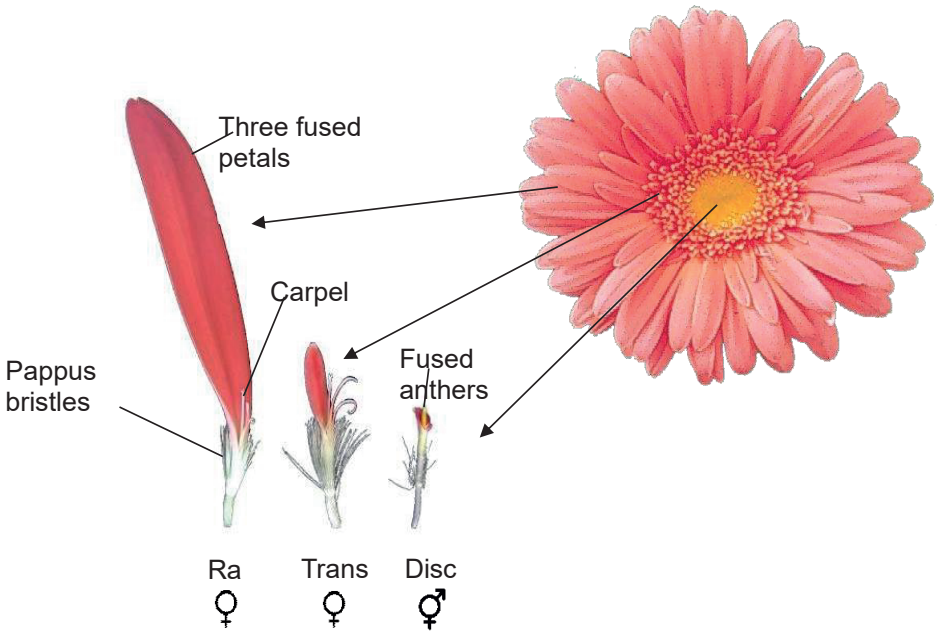

Figure 3. The structure of Gerbera inflorescence and different flower types. Ray, trans and disc flowers differ in size, morphology and sex. Ray and trans flowers are female, disc flowers are hermaphrodite.

All flower types of Gerbera contain a calyx structure, which typical to Asteraceae, is not leaf-like. Usually in the developing eudicot flowers, the sepals form a cover protecting the emerging bud of the flower. In Gerbera protection is provided by involucral bracts, which surround the capitulum and cover the emerging inflorescence during the early stages of development. The calyx of Gerbera has acquired a function in seed dispersal by being attached to seeds and developing into a feather-like

structure, the pappus hairs, which indeed are modified sepals (Kotilainen *et al*. 2000). With later results, Zhang *et al*. (2017) speculate pappus identity to have originated from involucral bracts.

At early developmental stages all Gerbera flower types have five petal primordia in their corolla. They fuse to form a typical ring primordium. As development proceeds, three of these primordia remain fused to form the petal ligule. This ligule is conspicuous in the ray and trans flowers, but not remarkably long in the disc flowers. The two abaxial petals remain small and rudimentary in all flower types. The stamen primordia in all floral types start to develop normally at the early stages, but their development is soon arrested in the ray and trans flowers, where they form only staminodes in the mature flowers (Kotilainen *et al.* 2000). In the disc flowers where stamens develop until maturity, they fuse postgenitally to enclose the carpel. Carpel development is identical in all Gerbera flower types. The gynoecium consists of two fused carpels. The Gerbera ovary is positioned inferior in contrast to the common model species (Arabidopsis, Antirrhinum and Petunia) and carries a single ovule.

1.11.1 Gerbera MADS domain genes and ABC model

In Gerbera, *GSQUA1* was reported previously (Yu *et al*. 1999), but functional data was not available. Since then, several Gerbera *SQUA*-like genes were either cloned by PCR or recognized from Gerbera EST collections (Laitinen *et al*. 2005).

The Gerbera B class genes are *GGLO1* and *GDEF2*, and they are strongly expressed in petals and stamens as the ABC model predicts. The transgenic Gerbera lines show phenotypic changes, which are consistent with the B class gene action (Yu *et al*. 1999). Another Gerbera B class gene, *GDEF1* belongs to TM6 lineage, which seems to be lost from both Arabidopsis and Antirrhinum. Gerbera also harbours *GDEF3*, which is paralogous gene to *GDEF2*. This paralogous gene pair possibly originates from a duplication event in Asteraceae. *GDEF1* obviously has no function in petal development, but has a redundant role in stamen development (Broholm *et al.* 2010)

The expression pattern of Gerbera C class genes *GAGA1* and *GAGA2* matches the expression of Arabidopsis and Petunia C class genes. The phenotypic changes at organ level in transformant lines are consistent with the ABC model (Yu *et al*. 1999). While an A function comparable to Arabidopsis does not exist in Gerbera (I), it remains unclear what factors restrict the C function in the perianth.

No D class genes have been reported in Gerbera.

Gerbera *GRCD1* and *GRCD2* were reported as the whorl specific E function genes (Kotilainen *et al.* 2000, Uimari *et al.* 2004). Later *GRCD4* and *GRCD5* were identified from the EST collection (Laitinen *et al.* 2005). The current total number of SEP-like genes in Gerbera is eight (Elomaa *et al.* 2018). New single and double transgenic lines for *GRCD4* and *GRCD5* created with RNAi constructs produced recurrent phenotypes. Based on the phenotypes and gene expression data Zhang *et al.* (2017) showed *GRCD4* and *GRCD5* to be differently specialized in petal development.

Current high number of *AP1*- and *FUL/paleoAP1*-like genes in Gerbera most likely makes them at least partially redundant. The large number and the differential action of these genes might contribute to the complexity of Gerbera inflorescence. Redundancy is common in the MADS box gene family (Liljegren *et al*. 2000, Pelaz *et al*. 2000, Pinyopich *et al*. 2003, Vandenbussche *et al*. 2004). Redundant copies of the genes can be maintained in the genome as their functional redundancy increases mutational robustness and evolvability, possibly leading to evolutionary novelties (Wagner 2008).

In addition to *GSQUA2* described here, Gerbera has five other *AP1*- and *FUL*-like genes. *GSQUA3* has the *euAP1*- motif and based on phylogenetic analysis *GSQUA1* also appears to be related to *euAP1*-genes (Yu *et al*. 1999). GSQUA4, GSQUA5 and GSQUA6 all possess paleoAP1/FUL-motif. Some of these genes are likely to be the outcome of the gene duplication and have been maintained in the Gerbera genome as subfunctionalization has occurred.

High degree of redundancy between related genes makes it difficult to assess function to the genes of interest. Transgenic downregulation lines are especially problematic since only subtle changes, or no visible phenotype at all is observed when a related gene is substituting the function of the downregulated gene completely, or partially, leading to underestimation of the gene's function. Sometimes closely related genes are inadvertently downregulated together with the gene of interest, and the resulting phenotype is the outcome of several genes downregulated (Angenent *et al*. 1994, Ferrario *et al.* 2003). This might lead to the overestimation of the effect of a single gene. Overexpression lines are usually more informative but overproducing a certain protein may not have any effect, if it is not a limiting factor.

Our group has shown Gerbera B and C functions to behave according to the ABC model (Coen & Meyerowitz 1991, Yu *et al*. 1999, Broholm *et al.* 2010). Gerbera E function genes *GRCD1* and *GRCD2* were shown to be whorl specific (Kotilainen *et al*. 200, Uimari *et al.* 2004). We studied closely related E function genes *GRCD4* and *GRCD5* to see if whorl specificity applies to them (II). A function comparable to Arabidopsis (Irish & Sussex 1990, Mandel & Yanofsky 1995b) does not exist in Gerbera, but Gerbera contains at least six *GSQUA* genes (Yu *et al.* 1999, Laitinen *et al.* 2005, I) .

The function of these genes in Gerbera was not known and *GSQUA2* was studied further (I). Arabidopsis *SOC1* is a MADS box gene that has function as flowering time integrator (reviewed by Lee & Lee 2010). *Gh-SOC1* was identified from Gerbera EST collection (Laitinen *et al.* 2005) and was subjected to reverse genetics studies to deduct its role in Gerbera. The ABCDE model with Arabidopsis and Gerbera floral organ determination is presented in figure 4.

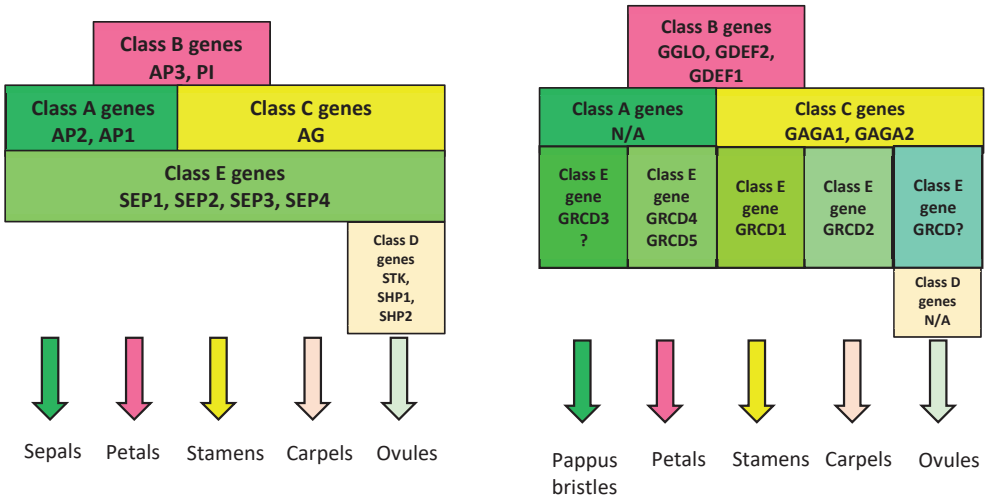

Figure 4. ABCDE model and floral development genes in Aradidopsis and Gerbera, Models according to Theissen *et al.* 2016 and Zhang *et al.* 2017.

2 AIMS OF THE STUDY

The purpose of this study was to identify genes and proteins putatively involved in *Gerbera hybrida* inflorescence development, and to particularly concentrate on the roles of the Gerbera MADS box genes outside the well-described B and C function genes. The findings presented here add to our previous knowledge of the floral development in Asteraceae.

The approaches used to study the Gerbera floral development in this thesis were candidate gene approach (I, III) and protein-protein interaction studies (II). In the candidate gene approach, reverse genetics method was utilized. MADS box genes that appeared interesting based on their sequence similarity to well-known genes, were transformed into Gerbera and the resulting phenotype of the altered gene expression was observed. Based on the phenotypic changes, and other information (e.g. expression patterns), the function of the gene could be deducted. This was applied to study the functions of Gerbera *SQUAMOSA*-like gene (I), and *Gh-SOC1* (III). Large-scale protein-protein interactions studies were conducted for prediction of Gerbera E class gene functions (II).

3 MATERIALS AND METHODS

Table 2 lists the materials and methods used in these studies. The more detailed descriptions can be found in the respective articles (I, II, and III).

Table 2. List of materials and methods used in studies presented in this thesis.

Yeast-two- and three hybrid method	I, II, III
DNA and RNA gel blots and hybridization	I, II, III
Isolation of plant DNA and RNA	I, II, III
In situ hybridization assay	I, II, III
Agrobacterium-mediated Gerbera transformation	I, II, III
PCR	I, II, III
(Scanning electron microscopy (SEM))	I, III
(Phylogenetic analysis)	I, II, III
(Organ and cell measurements)	III
Statistical analysis (T-test)	I
Tobacco protoplast isolation and electroporation	II
Gateway primer design and recombination reactions	I, II, III

The methods shown in parenthesis were used by my co-authors.

4 RESULTS AND DISCUSSION

4.1. Gerbera *GSQUA*-like genes and their sequence properties (I)

As discussed earlier, the existence of A function in plant model species is unclear (1.3.2). Our interest was to identify candidate genes homologous to the A class genes in Gerbera, and to conduct functional studies to reveal their role in Gerbera floral and/or inflorescence development.

To date, six *euAP1*- and *FUL*-like genes named as *GSQUA*s are known from Gerbera (Yu *et al.* 1999, Laitinen *et al.* 2005, I). Three *GSQUA* genes, *GSQUA2*, *GSQUA3*, and *GSQUA4* (I) were cloned by PCR using degenerate primers designed for the intervening sequence between conserved MADS and K domains utilizing an alignment containing MADS box genes from Gerbera and several other plant species (Purugganan *et al*. 1995, Yu *et al*. 1999). *GSQUA1* was isolated earlier by Yu *et al.* (1999) using MADS domain specific primer. *GSQUA5* and *GSQUA6* were isolated from Gerbera EST collection (Laitinen *et al*. 2005). *GSQUA5* was recovered as a full-length clone, but *GSQUA6* was missing sequence at the 5' end of the gene.

The relationships among *euAP1*- and *FUL*-like genes were clarified by the discovery of C terminal protein motifs by two groups simultaneously (Litt & Irish 2003, Vandenbussche *et al.* 2003b). EuAP1-like proteins possess two motifs, a transcription activation domain (Litt & Irish 2003) and a farnesylation signal (Galichet & Gruissem 2003), which is not present in all euAP1-like proteins. The function of these motifs remains unclear, but newer studies have questioned their impact on protein function diversification (Pabón-Mora *et al*. 2012). FUL-like proteins, instead, possess a hydrophobic motif, which is not found in the euAP1-like proteins.

GSQUA3 harbours C terminal protein motifs that place it among the euAP1-like proteins. *GSQUA1* sequence has a stop codon before the motifs, but it is classified as an euAP1-like sequence according to the phylogenetic analysis. The transcriptional activation domain of GSQUA3 is slightly varied (RSNELDLSL) from the consensus sequence (RRNaLaLT/NLa), but it is readily recognizable. The farnesylation signal of GSQUA3 (CFPS) differs from the most common consensus sequence (CFAA/T) found among the euAP1-like proteins in many plant species (Litt & Irish 2003). It is not clear how (or if) these variations affect the function of GQUA3. Comparison of the farnesylation signals among the closely related Asteraceae (sunflower *Helianthus annuus* and Chrysanthemum *Dendtranthema grandiflorum*) euAP1-like proteins shows them to be identical (Shchennikova *et al*. 2004, Shulga *et al.* 2008). The sequence CFPS might thus be Asteraceae specific, but still it remains within the definition of the CaaX for the signal (Hancock *et al*. 1991).

The results of the phylogenetic analysis of *GSQUA* genes confirmed the relationships suggested by the existence of the C terminal protein motifs. *GSQUA2* might be orthologous to *DEFH28* of Antirrhinum based on the maximum likelihood tree. In Antirrhinum, *DEFH28* regulates floral meristem identity and fruit development (Műller *et al*. 2001). At the time *DEFH28* function was published, it was speculated to be an ortholog of the Arabidopsis *FUL*. However later, the discovery of *AmFUL* proved this assumption wrong. *AmFUL* is the real ortholog of the Arabidopsis *FUL*, based on existence of the C terminal protein motifs and the phylogenetic analysis (Litt & Irish 2003). With *DEFH28*, *GSQUA2* shares the early flowering feature when ectopically expressed. The role in the fruit development cannot be compared, as it was not studied regarding *GSQUA2*.

FUL-like genes have been further divided into two groups, *euFUL* and *euFULII* (Litt 2007, Litt & Irish 2003, Litt & Kramer 2009). Arabidopsis *FUL* and Antirrhinum *AmFUL* (Gu *et al.* 1998, Litt & Irish 2003) belong to the first group, and Arabidopsis *AGL79* and Antirrhinum *DEFH28* (Műller *et al.* 2001, Litt & Irish 2003) to the latter group. Phylogenetic analysis shows Gerbera *GSQUA2* to belong to the *euFULII* group, while the other *FUL*-like *GSQUA* genes are closer to the *euFUL* group but form a Gerbera specific subclade. *GSQUA1* (Yu *et al*. 1999) and *GSQUA3* seem to be a recent paralogous pair of genes and based on the phylogenetic tree, they are co-orthologous to Antirrhinum *SQUA* (Huijser *et al*. 1992) after which they were named. The remaining *GSQUAs,* excluding *GSQUA2,* are co-orthologous to Arabidopsis *FUL* (Ferrándiz *et al*. 2000, Gu *et al.* 1998).

4.2 Expression domains of *GSQUA*-like genes

The existence and expression of *AP1/FUL*-like genes is not well characterized in Asteraceae species. Chrysanthemum *CMD111* groups together with *euAP1*-like genes (I) and has an expression pattern similar to *SQUA*-like genes. High expression was detected in the mature inflorescence bracts and petals of both flower types in sunflower (*Helianthus annuus*) (Shchennikova *et al*. 2004). The sunflower *HAM75* has been identified as a homologous gene to Arabidopsis *AP1* (Shulga *et al.* 2008). In the sunflower EST collection analysis a gene homologous to Arabidopsis *FUL* was identified (Blackman *et al.* 2011). Sunflower genome has been sequenced (Badouin *et al*. 2017), but no specific information about *FUL*-like genes has been published.

GSQUA1, GSQUA2 and *GSQUA5* genes were inflorescence abundant while *GSQUA*3, *GSQUA4* and *GSQUA6* were also expressed in leaves. None of the genes studied here were expressed in Gerbera roots. *GSQUA* genes, which were found to be expressed in Gerbera petals, were further analysed during different stages of ray flower petal development. Two of the analysed *GSQUA* genes had varying expression during petal development. *GSQUA2* was clearly more active in the early stages of petal development, whereas *GSQUA6* showed the opposite pattern, and was expressed

at higher level at the late stages of petal development. *GSQUA3* was found out to be expressed uniformly during petal development, as did *GSQUA5*, whose expression, however, seemed to intensify at the late stages 10 and 11. *GSQUA4* was poorly expressed in Gerbera petals.

To get a more detailed analysis of gene expression at different stages of inflorescence development, sections from different sizes of Gerbera capitula (6-17 mm) were subjected to *in situ* hybridization assay. The expression pattern of different *GSQUA* genes was widespread. Only *GSQUA4* showed a more restricted pattern in stamens and the tip of the carpel. All other GSQUA genes were widely expressed in most floral organs. However, all *GSQUA* genes showed specific expression patterns despite being closely related. Vasculature expression was the only common feature shared by all *GSQUA*s. The vasculature of Gerbera receptacle and petals was the only location where *GSQUA1* expression was visible. Vasculature expression has been detected also for *FUL* (Gu *et al*. 1998), but which is not a unifying factor for the *euAP1*, *euFUL* and *euFULII* genes. This expression might reflect a function in the developing vasculature, but this hypothesis has not been studied further, and thus its functional significance is not clear. *GSQUA2* and *GSQUA5* shared a wide expression pattern, including all floral parts. The expression pattern of *GSQUA3* and *GSQUA6* was narrower at floral organ level, and *GSQUA4* expression was confined to the reproductive organs. A more detailed examination of developing ray flowers revealed *GSQUA2* expression to be localized on the adaxial surface of the petal. The strong expression of *GSQUA2* marked the location of emerging flowers in the yet undifferentiated capitulum. None of the expression patterns of *GSQUA* genes, however, matches that of either Arabidopsis *AP1* or *SQUA* of Antirrhinum by being abundant in the sepals and petals at the early stages of floral development (Mandel *et al*. 1992, Huijser *et al.* 1992).

Orthologous genes in Antirrhinum (*SQUA*, *DEFH28* and *AmFUL*) show partially overlapping expression patterns with Gerbera *GSQUA* genes. Antirrhinum genes are expressed in petals and gynoecium, *SQUA* and *AmFUL* also in sepals. Expression in stamen is very weak (Preston & Hileman 2010). Also *GSQUAs* expression is typically seen in petals, carpel and ovary. *GSQUA* expression can be detected also in stamen to a varying degree. Only *GSQUA1* shows a very limited expression pattern confining to vascularature and petals (Yu *et al.* 1999).

The wide expression pattern of *GSQUA2* resembles more the expression pattern of *FUL* than that of *AP1*. *FUL* expression has been detected in various floral organs, meristems and vegetative tissues as well as in all floral whorls in Arabidopsis (Mandel & Yanofsky 1995a, Gu *et al.* 1998). *FUL* is expressed at the early stages of floral development when floral meristem identity is determined, and again later during silique development. The hypothetical role of *GSQUA2* in Gerbera achene development is hard to analyse due to morphological differences. Gerbera does not bear a fruit similar to Arabidopsis but has an inferior ovary and the fruit (achenes) are indehiscent. The putative

late role of *GSQUA2* in Gerbera fruit development might be entirely different than in Arabidopsis or lacking completely. Despite of *GSQUA2* being highly expressed in the ovary inner wall and in the ovule in Gerbera, no homeotic changes in overexpression lines were discovered in the ovary and the ovule at stage 8. As the most dramatic changes in overexpression lines of *FUL* appeared at the cell type level, this cannot be ruled out in Gerbera, which was not studied at the level of cellular differentiation. All this data taken together, *GSQUA2* does not provide the A function in Gerbera. In the light of the expression patterns for other *GSQUA*s, none of them supports the existence of the A function in Gerbera either. As described in the introduction (1.1.1) the existence of A function outside Arabidopsis is debated.

The number of Gerbera GSQUA genes (currently six) is high compared to Arabidopsis and Antirrhinum, which both have three of *AP1*- and *FUL*-like genes (*AP1*, *CAL*, and *FUL*, *SQUA*, *DEFH28*, and *AmFUL*) (reviewed in Parenicová *et al*. 2003 and Schwarz-Sommer *et al.* 2003). This might also reflect the whole genome duplication history of Asteraceae. Several WGD events have taken place in Asteraceae (Barker *et al.* 2008, Barker *et al*. 2016, Huang *et al*. 2016). An early Asteraceae WGD possibly affected the genes that led to the development of complex capitulum (Chapman *et al*. 2008, Chapman *et al.* 2012). This Asteraceae specific WGD has also been linked to the large *CYC/TB1* gene family in Gerbera and the gene family's potential role in complex capitulum development (Broholm *et al*. 2008, Tähtiharju *et al*. 2012). Sunflower is closely related to Gerbera but only two *SQUA*-like genes have been identified to date (Shulga *et al.* 2008). The whole sunflower genome has been published (Badouin *et al.* 2017). Another closely related Asteraceae species, Chrysanthemum has been reported to contain three *AP1*- and *FUL*-like genes (Shchennikova *et al*. 2004). The genome of Chrysanthemum (*Chrysanthemum seticuspe*) was recently assembled by Hirakawa *et al.* (2019).

The high number of *GSQUA* genes leads to speculation of these genes being associated with the complex inflorescence structure of Gerbera. The evolutionary origin of these genes is also interesting. It seems that more copies of duplicated genes were retained in Gerbera genome. It would be interesting to extract all *euAP1*- and *FUL*/*paleoAP1*-like genes from other Asteraceae species and compare their numbers and inflorescence complexity to *Gerbera hybrida* to see whether any correlation between the two traits exist.

4.3 Overexpression of *GSQUA2* – life of a rock star: live fast, die young

GSQUA2, *GSQUA3* and *GSQUA5* were transformed into Gerbera to investigate their functions. Of these, only *GSQUA2* lent itself for functional analysis. Overexpression or downregulation of *GSQUA3* or *GSQUA5* did not produce a consistent phenotype for analysis. *Antisense* RNA for *GSQUA2* was produced, but we were not able to confirm downregulation of the endogenous gene. Strong overexpression of *GSQUA2* under 35S promoter consistently produced an early flowering, dwarf phenotype with vegetative abnormalities. Flowering was accelerated by approximately four months as the plants overexpressing *GSQUA2* flowered after two months in the greenhouse. Wild type Gerbera 'Terra Regina' typically takes about six months to flower, depending on the season. Flowering time of the *35S-GSQUA2* plants was not systemically recorded, but the time frame is based on the observation in the greenhouse. The plants were tiny compared to the wild type plants, typically only 20 cm tall. At the inflorescence level, the number of flowers in the overexpression lines was dramatically reduced. Instead of the normal ca. 900 flowers, the transformant inflorescences contained only 500 flowers. Vegetative abnormalities were seen in leaves, which curved strongly adaxially. Another vegetative abnormality was linked to the loosening of the leaf rosette typical to Gerbera. The growth habit of Gerbera is sympodial, with extremely short internodes that creating a tightly packed leaf rosette. The leaf internodes were elongated in the overexpression lines, and the resulting small plants were wobbly.

The function of *GSQUA2* was linked to flowering time. In Arabidopsis, overexpression lines of FUL flowered early and formed clearly reduced number of rosette and cauline leaves (Ferrándiz *et al*. 2000, Melzer *et al.* 2008, Balanzà *et al*. 2014). Despite that early flowering is quite a common phenomenon associated with the MADS box gene overexpression, also associated with genes that are not directly related to flowering time (Lee *et al.* 2000, Elo *et al.* 2001, Ferrario *et al.* 2004, Fornara *et al*. 2004), this is the only case when accelerated flowering in Gerbera has been observed. In Petunia, *AP1/SQUA* subfamily contains four members (*PFG*, *FBP26*, *FBP29* and *euAP1*) (Immink *et al.* 2003, Morel *et al.* 2019). They are largely redundant and required for inflorescence meristem identity. Suppression all the subclade members in Petunia produced fully functional terminal flowers. Floral meristem identity is not dependent on AP1/SQUA subclade genes in Petunia, but the authors were not able to exclude residual gene action. Petunia *AP1/SQUA*-like genes act as B function suppressors in first whorl together with *BEN*/*ROB* genes, and were not involved in Petunia petal development and thus cannot be classified as A function genes comparable to Arabidopsis (Coen & Meyerowitz 1991). Morel *et al.* (2017) proposed a floral organ identity model for Petunia in which the original A function is replaced by boundary setting mechanism that patterns B and C functions (Morel *et al.* 2017).

Antisense lines of *GSQUA2* did ot show any effect on the vegetative, or the floral phenotype. The closely related genes might function redundantly and substitute the function of *GSQUA2*. The effect of *GSQUA2* downregulation was only manifested in flowering time, which was considerably delayed and correlated with the overexpression phenotype. MADS box proteins function by forming higher order complexes (Egea-Cortines *et al.* 1999, Honma & Goto 2001). With the interaction capacity shown by yeast two and three hybrid assays (II), GSQUA2 is likely to regulate multiple target genes. The effect of *FUL*-like genes on flowering time is well described (Melzer *et al.* 2008, Fornara *et al.* 2004), and dwarfism is reported in many cases to be related to overexpression of MADS box genes (e.g. Sung *et al*. 1999, Jeon *et al*. 2000, Jang *et al*. 2002, Wang *et al*. 2004). The release of leaf rosette and the leaf internode elongation has not been reported earlier with any MADS box gene overexpression in any other plant species. Typically, Gerbera stays vegetative until 7-26 rosette leaves are formed, then becomes reproductive and forms capitula. The sympodial growth progresses by the axillary bud that is located at the axil of the penultimate leaf becoming reactivated. Excess amount of GSQUA2 interferes with this growth habit and leads to elongated leaf internodes in overexpression lines. Either the GSQUA2 protein releases an arrest on stem growth forming between the rosette leaves, or its activity cannot be repressed due to ectopic expression.

In addition to its role in flowering time and fruit development in Arabidopsis, *FUL* has an important function in plant growth and architecture (Bemer *et al.* 2017). In Arabidopsis, *FUL* was found to activate several genes that regulate stem growth and architecture. Genes involved in cytokinin and auxin signaling and light sensitive growth of hypocotyls were identified. The authors concluded that *FUL* regulates a network of genes that have opposite functions in stem growth. They also concluded that other factors such as hormone concentration and light quality contribute to the regulation. These results are interesting in respect of the 35S:SQUA2 lines with elongated rosette leaf internodes. Small Auxin Upregulated RNA 10 (*SAUR10*) is a growth regulator that, in Arabidopsis, is induced by different hormones and light conditions and is also a direct target of *FUL*. In normal conditions *FUL* represses SAUR10, but in *35S:FUL* plants branch angles are larger (Bemer *et al*. 2017). Similar additional growth regulators could contribute to internode elongation in Gerbera rosette leaves.

Unfortunately, the functions of the other *GSQUAs* are still unclear. Some degree of redundancy must exist among the *GSQUA* genes as single gene downregulation did not produce any floral phenotype. Knocking down several *GSQUA*s simultaneously should give a better idea of their functions in Gerbera.

4.4 Interaction profile of Gerbera GSQUAs

Two Gerbera euAP1- and two FUL-like (GSQUA1, GSQUA2, GSQUA3 and GSQUA5) proteins were included in the pairwise yeast assays. No heterodimer formation was detected among these proteins comparable to Petunia (Immink *et al.* 2003), but the FUL-like GSQUA5 reacted with the SEP-like GRCD5 and the FUL-like GSQUA2 interacted with the SEP-like GRCD2. The interactions among SEP3/4 and AP1/FUL genes have been detected in Arabidopsis (Smaczniak *et al.* 2012). Interestingly, GSQUA5 remained inactive in pairwise assays but in yeast three hybrid assay GSQUA5 was active in several trimeric complexes displaying both strong and weak interactions with GRCDs and Gerbera C function proteins. More limited interactions were also detected among GSQUA5, GRCDs and B function GDEF1 and GDEF2. GSQUAs are not A function proteins, and thus are not involved in sepal identity determination in Gerbera, but rather are involved in floral meristem identity determination and floral meristem transition.

AP1-like proteins have been observed to connect proteins as bridging proteins, and to facilitate the formation of protein quartets (Kaufmann *et al.* 2005, Theissen & Melzer 2007). For Gerbera, though, this hypothesis does not seem to be applicable, since the interaction capacity of the GSQUA proteins is limited (II). All petunia FUL-like proteins (FBP29, PFG, and FBP26) and the Arabidopsis FUL behave according to the hypothesis and interact with a wide range of MADS domain proteins (Immink *et al*. 1999, Immink *et al*. 2003, de Folter *et al*. 2005, Immink *et al.* 2009). GSQUA2 was the most active GSQUA protein and interacted with three different Gerbera MADS domain protein partners. The EuAP1-like GSQUA1 and GSQUA3 both formed two interactions with the SEP-like GRCD proteins. GSQUA5 did not interact with any of the tested MADS domain proteins. Interaction between Arabidopsis FUL and SEP4 protein have been identified in *in planta* studies by Smaczniak *et al.* (2012). They also identified SOC1 as major interaction partner for FUL. Gh-SOC1 interaction was detected with SEP-like GRCD1 and GRCD5 and FUL-like GSQUA3 (III). These interactions are comparable to those of Arabidopsis SOC1 that support a FUL/SOC1 protein complex existence in floral transition (Smaczniak *et al*. 2012). Immink *et al.* (2012) identified multiple MADS domain and other protein dimers regulating SOC1 by binding to its regulatory sequences upstream and downstream.

The interaction among GSQUA2 and GRCD2 was the most interesting one since the SEP-like GRCD2 has a function in determining carpel and floral meristem identity as well as inflorescence determinacy (Uimari *et al*. 2004). GSQUA2/GRCD2 dimer also gave rise to strong autoactivation. This was dependent of dimerization as neither of the proteins alone was able to activate transcription in yeast. This strong activation reaction might reflect the importance of this dimer in Gerbera

inflorescence development. As the temporal and spatial expression of both genes coincide, the interaction observed in yeast may occur also *in planta*.

Whilst the GSQUA proteins remained rather inactive in the pairwise assays, they were more active in higher order complex formation. The B function dimer GGLO1/GDEF2 was able to recruit all GSQUA proteins into a threesome complex. Similarly, when combined with a SEP-like protein (GRCD) and a C function protein (GAGA), all GSQUA proteins reacted by forming higher order complexes. None of the GSQUA proteins formed dimers between themselves in pairwise assays, but addition of a SEP-like GRCD protein stabilized some complexes containing two different GSQUA proteins. SEP-like GRCD4 and GRCD5 were able form ternary complexes among many Gerbera MADS domain proteins that did not react in pairwise assays (II).

The GSQUA protein interaction with the SEP-like GRCDs (GRCD4, GRCD5) and with the B function dimer (GGLO1/GDEF2) is consistent with an A function according to the ABC model, but the interaction with the C function proteins (GAGA1 or GAGA2 + GRCD in ternary complex) certainly contradicts that. The A and C function are supposed to be mutually antagonistic according to the ABC model (Coen & Meyerowitz 1992), and in Arabidopsis *AP1* is not expressed simultaneously with the *AG* in the same cells (Gustafson-Brown *et al*. 1994). AP1 is not capable of repressing AG alone in sepals and petals, but it recruits the non-MADS domain proteins LEUNIG and SEUSS (Gregis *et al*. 2006, Sridhar *et al.* 2006) for a complex together with the MADS domain proteins AGL24 and SVP (Gregis *et al*. 2006). Besides the floral homeotic function, AG also controls the meristematic state of flower primordia and downregulates the meristematic organizing gene WUCSHEL together with unidentified factors (Lohmann *et al*. 2001, Lenhard *et al.* 2001). In petunia, these factors are identified as MADS domain proteins (Ferrario *et al.* 2006).

Considering *AP1* as a flowering time gene, the interaction with C and E function genes has a different implication. In the large-scale analysis of Arabidopsis MADS domain protein interactions, an interaction of AP1, SEP3 and AG was discovered (Immink *et al.* 2009). The authors speculate this complex of flowering time and floral organ proteins to function in negative auto-regulatory feedback mechanism. In floral organ primordia the function of floral inducing MADS box genes is downregulated. The floral inducing MADS domain proteins and the floral organ MADS domain proteins dimerize and form a negative autoregulatory loop (de Folter *et al*. 2005). AP1 has been shown to downregulate the flowering time genes SOC1, SVP and AGL24 (Liu *et al*. 2007).

4.5 Gerbera SEP-like genes: expression patterns and transgenic phenotypes (II)

Arabidopsis harbours four *SEP* genes (*SEP1/2/3/4*) that are functionally redundant (Honma & Goto 2001). In phylogenetic analysis the angiosperm *SEP*-like genes form two subclades*, SEP1/2/4* and *SEP3*. Gerbera contains eight *SEP*-like genes, called *GRCD*s (Gerbera Regulator of Capitulum Development) (Kotilainen *et al*. 2000, Uimari *et al*. 2004, II, Zhang *et al.* 2017, Elomaa *et al.* 2018). Phylogenetic analyses performed by Zhang *et al*. (2017) and in II gave similar results regarding the *GRCD* genes. *GRCD4* grouped in *SEP4* subclade and *GRCD5* in *SEP3* subclade. *GRCD1* formed an Asteraceae specific subclade together with Helianthus and Chrysanthemum sequences. This subclade was also present in the phylogenetic analysis performed in II. The new analysis discovered recent duplications in *SEP1*/2 (*GCRD2* and *GRCD7*) and *SEP3* (*GRCD5* and *GRCD8*) subclades. *GRCD1* and *GRCD2* are well characterized by Kotilainen *et al*. (2000) and Uimari *et al*. (2004). *GRCD3* was first discovered by Kotilainen *et al.* (2000) and later found from the EST collection (Laitinen *et al.* 2005). Both *GRCD4* and *GRCD5* were isolated from the Gerbera EST collection (Laitinen *et al*. 2005).

The expression of *GRCD3*, *GRCD4* and *GRCD5* was analysed by RNA gel blots and *in situ* hybridization analysis. The strongest expression of *GRCD3* was seen in the inflorescence, petals and ovary by the RNA blots. Weaker expression was also detected in the carpel, receptacle, stamens, pappus bristles and bracts. Timewise, the expression was strong in the earlier stages of ray flower petal development (stages 1-7, for stages see Helariutta *et al*. 1993) and the signal was barely detected at the final stages studied. When analysed in more detail, *in situ* hybridization analysis revealed *GRCD3* expression in several locations. The ovule, carpel and petals gave a strong signal, whilst a weaker expression was detected in stamens and pappus bristles. *GRCD4* was strongly expressed in young Gerbera inflorescences (diameter 6-16 mm), petals and ovary. Weaker signal for expression was detected in the carpels, pappus bristles, receptacle and stamens in RNA gel blots. *GRCD4* was expressed throughout the Gerbera ray flower petal development, only slightly fading towards the later developmental stages. *In situ* hybridization analysis of *GRCD4* confirmed the results of RNA gel blot analysis by showing that *GRCD4* is widely expressed in all floral organs. *GRCD5* expression was floral specific, and the strongest signal was detected in young inflorescences (diameter 6-16 mm) and petals. Weaker expression was visible in other floral whorls (stamen, carpel, ovary) and receptacle. Very low level of expression was seen in the bracts and pappus bristles. In contrast to *GRCD3* and *GRCD4*, which were expressed at the earlier stages of ray flower petal development, *GRCD5* showed clear upregulation towards the later stages of the petal development. *In situ* hybridization results for *GRCD5* were comparable to those of *GRCD4* by showing a widespread expression in all the floral organs. In this thesis (II), neither overexpression

nor downregulation with sense or antisense constructs of *GRCD4* or *GRCD5* alone produced lines with consistent phenotypic changes in floral or vegetative features.

Zhang *et al.* (2017) further characterized *GRCD4* and *GRCD5*. In phylogenetic analysis, additional gene family members *GRCD6*, *GRCD7* and *GRCD8* were included. The expression patterns of GRCD 4 and GRCD5 during capitulum and flower type development were determined by quantitative real-time PCR. The results were contradictory to the earlier results obtained by probing RNA blots with gene specific probes (II). *GRCD4* expression was not detected in inflorescence meristem but was found in developing flower primordia. Both *GRCD4* and *GRCD5* were localized in the emerging flower primordia. The highest expression for *GRCD4* and *GRCD5* was seen in petals and pappus bristles.
GRCD4 expression was detected in ovary walls and *GRCD5* in developing ovule. Zhang *et al.* performed *insitu* hybridization analysis by using shorter gene specific probes compared to II.

Single and double transformant lines for *GRCD4* and *GRCD5* were obtained by using RNAi constructs (Zhang *et al*. 2017). The single transformant lines displayed mild phenotypes compared to the most severe double transformant lines for *GRCD4* and *GRCD5*, which showed a strong phenotype as predicted by the results in II. These transgenic lines showed the hypothesis of GRCD4 and *GRCD5* encoding the general E function in Gerbera that was speculated based on their PPI behaviour not to be entirely true (II). Zhang *et al.* (2017) showed *GRCD4* and *GRCD5* to be involved in Gerbera petal development. *GRCD4* controls petal cell specification and ovary wall development, whilst *GRCD5* is involved in petal cell proliferation and expansion.

4.6 Gerbera MADS domain proteins PPI: dimers and higher order complexes

All available Gerbera MADS domain proteins were tested for pairwise interactions by yeast two-hybrid method. MADS domain proteins are thought to exert their function by forming tetrameric complexes, which bind to their target sequences (Theissen 2001, Theissen & Saedler 2001). We used yeast three-hybrid system to assay whether the introduction of a third MADS domain protein into a pair of previously non-reacting proteins results in interaction. The SEP-like proteins are known for their capacity to recruit inactive proteins into trimeric complexes (Honma & Goto 2001, Immink *et al.* 2009). Out of 531 possible combinations 313 were tested. The tested Gerbera MADS domain protein pairwise interactions are presented in figure 5.

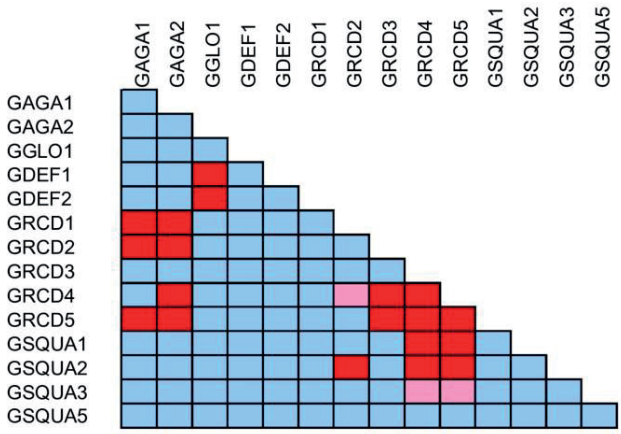

Figure 5. Yeast two-hybrid analysis of protein-protein interactions among Gerbera MADS domain proteins. Red, strong interaction; pink, weak interaction; blue, no interaction detected (II, Ruokolainen *et al.* 2010, BMC Plant Biology 10: 129).

Results of yeast three-hybrid analysis of ternary protein complex formation among Gerbera MADS domain proteins are shown in figure 6.

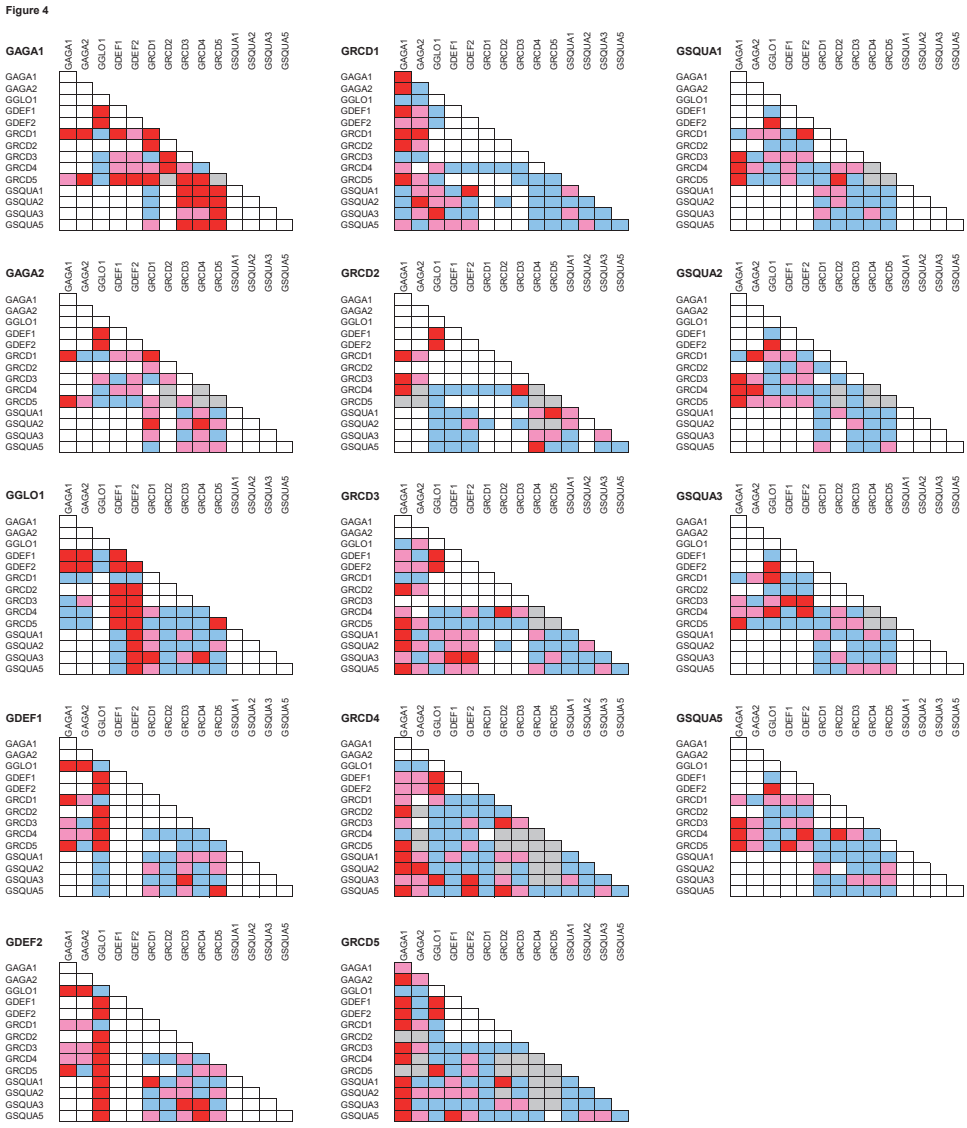

Figure 6. Yeast three-hybrid analysis of ternary protein complex formation among Gerbera MADS domain proteins. Red, strong interaction; pink, weak interaction; blue, no interaction; grey, interaction test uninformative; white, not tested (II, Ruokolainen *et al.* 2010, BMC Plant Biology 10: 129).

Some *in vitro* data has been shown for MADS domain protein tetramers (Melzer *et al*. 2009, Melzer & Theissen 2009), but *in planta* evidence is still scarce. The Gerbera MADS domain protein higher order complex formation data raises new questions of whether the complexes could contain more

than the earlier proposed four proteins, or if they could be transient in nature, different proteins participating at the different stages. Other authors have discussed the possibility of the complexes being larger than tetramers (Immink *et al.* 2009). The Gerbera MADS domain proteins show capacity for the both types of increased complexity. In affinity purification coupled with mass spectrometry, the bait protein is purified from a biological sample near physiological condition to keep the protein complex intact. The co-purified proteins are then identified by mass spectrometry (reviewed in Bontinck *et al*. 2018). This approach could shed more light to the true nature of MADS protein ternary complexes. Smaczniak *et al.* (2012) have shown MADS domain proteins to form large protein complexes that include also non-MADS transcription factors and chromatin remodelling proteins.

4.6.1 Gerbera B protein complexes and stamen identity regulation

The B function proteins are highly conserved, both developmentally and biochemically (reviewed in Irish & Kramer 1998). As Gerbera flower types show differentiation in petal size and in sex, the B function proteins (GGLO1, GDEF1, GDEF2) were of special interest. Obviously, lack of expression of the Gerbera B function proteins is not responsible for abortion of stamens in ray and trans flowers as they are strongly expressed during the process (Yu *et al.* 1999). The B function dimers GGLO1/GDEF2 and GGLO1/GDEF1 (GDEF1 is TM6-like) reacted with the C function proteins GAGA1 and GAGA2. The GGLO1/GDEF2 was accepted by all GSQUAs, but the GGLO1/GDEF1 pair was rejected. This further illustrates the difference between GDEF1 and GDEF2. GDEF1 is proposed to have a role in stamen development (Broholm *et al.* 2010). None of the E function proteins (GRCDs) interact with B function proteins in pairwise assays, but the B function dimers show strong interaction with all Gerbera E function proteins excluding GRCD1. In Gerbera, the interaction capacity of the dimer GDEF2/GGLO1 is higher than that of single B function proteins, but the single B function proteins also interact with the C function GAGA and the E function GRCD proteins. Single B function proteins were also involved in active complexes, in which none of the studied proteins interact in pairwise assays. The significance of these single B function protein interactions in the higher order complexes is not known.

Gerbera PPI results show that the B function dimers can recruit other MADS domain proteins into transcriptional complexes. Interestingly, the B function protein dimers have intrinsic activation capacity. When the B function dimers GGLO1/GDEF2 and GGLO1/GDEF1 were combined with an empty activation vector, or tested completely without an activation vector, growth still resulted on plates selecting for weaker interaction. Dimerization of the MADS domain proteins is necessary for their target sequence binding (CArG boxes) activity, but obviously in Gerbera the B function proteins become also transcriptionally active. Surprisingly, also single Gerbera B function proteins were participating in higher order complexes. Thereby, dimerization does not seem to be an obligatory

requirement for Gerbera B function proteins. Previously this has only been described in tomato, and typically the B function proteins participate in higher order complexes only as dimers (Leseberg *et al*. 2008). These ternary complexes were observed for different combinations with GRCDs and GSQUAs. GDEF1 and GDEF2 reacted with GAGAs and GRCDs, whilst GGLO1 remained an outsider in these trimeric interactions.

Based on the PPI studies of the Gerbera MADS domain proteins, we propose protein complexes which might be involved in Gerbera stamen development. Both the B function protein dimer GDEF2/GGLO1 and the C function protein (GAGA1 or GAGA2) are necessary. GDEF1 has been linked to stamen development (Broholm *et al*. 2010), and due to its expression pattern, the dimer GDEF1/GGLO1 might also be involved in Gerbera stamen development. GRCD1 is necessary for staminoid development in the marginal flowers and both Gerbera C function proteins (GAGA1, GAGA2) interact with GRCD1 (Kotilainen *et al*. 2000). While GRCD1 does not directly interact with the B function protein dimer, the GRCD1/GAGA1 or GRCD1/GAGA2 dimer could interact with B function proteins. GRCD1 forms a higher order complexes with GAGA1 and GDEF1 (II), also a weaker interaction is seen with GAGA1 and DGEF2. GAGA2 reacts weakly with GRCD1 and GDEF1/2. The development of functional, pollen producing stamens in disc flowers could be an outcome of GRCD1 being outcompeted by other GRCDs in the complex formation. Earlier microarray analysis supports this hypothesis by showing that *GRCD1* is upregulated in ray flowers compared to disc flowers (Laitinen *et al*. 2006).

4.6.2 Gerbera C protein complexes

According to ABCE model, C function proteins are involved in complexes regulating stamen and carpel identity together with B and E function proteins (reviewed in Theissen *et al*. 2016).
In pairwise assays, Gerbera C function proteins GAGA1 and GAGA2 interacted only with the SEP-like proteins GRCD1, GRCD2 and GRCD5, GAGA2 reacted also with GRCD4. Based on the function of GRCD4 and GRCD5 in petal development and low expression levels in reproductive organs according to Zhang *et al.* (2017), the interaction with C function GAGA proteins that are involved in reproductive organ development, is not according to the ABC model. Gerbera C function proteins, GAGA1 and GAGA2, showed differential behaviour in yeast three-hybrid studies. GAGA1 was a more active player in forming higher order complexes than GAGA2. These proteins also shared interaction partners. GAGA2 preferred GCRD1 and GSQUAs as partners, while GAGA1 was active with other GRCDs and GSQUAs. In Gerbera E function GRCD1 and GRCD2 are necessary for correct floral organ determination and mediating the C function activity by GAGA1 and GAGA2 (Kotilainen *et al*. 2000, Uimari *et al.* 2004).The interaction maps are found in publication II.

The interaction data published in II verifies the former findings that GRCD1 and GRCD2 are specific and prerequisite activators of the C function (GAGA1, GAGA2) in Gerbera. These genes are expressed in stamens and carpels, and they both interact with the C function proteins. However, they are not able to complement each other in the wrong whorl, e.g. GRCD1 cannot replace GRCD2 in defining the carpel identity, and vice versa for the stamen identity (Kotilainen *et al.* 2000, Uimari *et al*. 2004). Plants downregulated for *GRCD1* have petal-like organs in the place of the staminoides in the ray flowers. However, the developed stamens of the central disc flowers are not affected. This raises the possibility that an unidentified MADS domain protein is taking over the function of GRCD1 in the disc flowers (Kotilainen *et al.* 2000). Similarly, Arabidopsis AG has a dual role in the development, first having an early function in floral organ development, and later, as suggested to be involved in anther development (Ito *et al.* 2007).

4.6.3 Gerbera E protein complexes

The behaviour of GRCD (SEP-like) proteins in interaction studies using Y2H was remarkably different. The previously characterized GRCD1 (Asteraceae specific sister clade to SEP3), which affects stamen development, and the carpel-linked GRCD2 (SEP1/2 clade) (Kotilainen *et al.* 2000, Uimari *et al*. 2004) showed a very limited interaction capacity when tested with all available Gerbera MADS domain proteins. In contrast, the related GRCD4 and GRCD5 had numerous interaction partners. C terminal deletions were introduced to GRCD4 and GRCD5 due to their strong autoactivation character. Both GRCD4 and GRCD5 interacted with eight other Gerbera MADS domain proteins. Moreover, they were the only proteins forming homodimers in the set of tested proteins.

As previously shown, GRCD1 and GRCD2 interacted with the C function proteins GAGA1 and GAGA2 (Uimari *et al*. 2004). For GRCD1, no new interacting protein partners were found, while GRCD2 readily accepted GRCD4 and the FUL-like GSQUA2 as dimerization partners. GRCD2 is involved in Gerbera floral development and controls carpel and floral meristem identity as well as inflorescence determinacy (Uimari *et al.* 2004). Interaction between GRCD2 and the Gerbera C function genes GAGA1 and GAGA2 has been previously studied and the result was verified in this study. This interaction reflects GRCD2's role in controlling homeotic and floral meristem identity. Both dimers GRCD2/GAGA1 and GRCD2/GAGA2 possessed strong transcriptional activation capacity, which none of the proteins had alone. The dimer GRCD2/GSQUA2 showed similar transcriptional activation. GSQUA2 strongly induces the flowering when overexpressed (I). The expression patterns of both GRCD2 and GSQUA2 show strong expression at the early stages of the floral development in inflorescence meristem and their expression patterns overlap enabling the interaction *in planta*.

Gerbera SEP-like protein, GRCD3 showed limited interaction capacity compared to the other GRCDs, interacting in pairwise assays only with GRCD4 and GRCD5. Zhang *et al.* (2017) suggest GRCD3 to contribute to pappus and petal identity and showed that GRCD4 and GRCD5 are involved in determining Gerbera petal identity. In phylogenetic analysis GRCD3 groups to AGL6 clade (II, Zhang *et al.* 2017). Petunia AGL6-like gene *PhAGL6* is involved in petal and anther identity specification and shows redundant action with other Petunia SEP-like genes (Rijpkema *et al*. 2009). The expression patterns also indicate a role in ovary/ovule development. Petunia AGL6 interacted with C and E function and SOC1-like proteins (de Folter *et al.* 2005). An interaction comparable to Petunia for GRCD3 was seen with GRCD4 and GRCD5. Weak interaction was detected among GRCD3, Gh-SOC1 and Gerbera C function proteins GAGA1 and GAGA2 (III). The interaction behaviour of Petunia AGL6 and Gerbera GRCD3 is similar and reflects their close phylogenetic relationship.

The dramatically different behaviour of the closely related proteins, GRCD1, GRCD2, GRCD4, and GRCD5, clearly illustrates how no interpretations on the protein function should be drawn based on the sequence similarity, or even close relationship shown by a phylogenetic analysis. The lack of clear phenotypes in transgenic Gerbera lines for GRCD4, and GRCD5, further suggested special function of these proteins. Based on yeast two-hybrid assay results, the function of these proteins is most likely at least partially redundant, and thus a transgenic plant with both *GRCD4* and *GRCD5* activity suppressed should result in a phenotype with a severely disturbed inflorescence. The single and double transgenic lines of *GRCD4* and *GRCD5* with RNAi constructs created by Zhang *et al*. (2017) helped to deduct the functions of these genes. According to the PPI results *GRCD4* and *GRCD5* seemed to provide the general E function in Gerbera comparable to that of Arabidopsis *SEP1* and *SEP3* genes (Honma & Goto 1999, Pelaz *et al*. 2000, Ditta *et al*. 2004). Their large interaction capacity and wide expression pattern in all floral whorls supported this hypothesis. The general E function hypothesis was shown not be entirely correct by single and double transgenic RNAi lines suppressing *GRCD4* and *GRCD5* function. Both GRCD4 and GRCD5 were shown to be involved in Gerbera petal development (Zhang *et al.* 2017). Whilst their interaction map is not identical, some redundancy is likely to exist. This was supported by the lack of clear phenotype when either *GRCD4* or *GRCD5* was downregulated by *antisense* constructs (II). The single transgenic lines for GRCD4 and GRCD5 downregulated with RNAi constructs showed variable degree of phenotypic changes in petals (Zhang *et al.* 2017). However, no interaction studies were re-performed for these proteins by Zhang *et al.* (2017). Despite their specific roles in petal development, both proteins GRCD4 and GRCD5 harbour transcriptional activation domain and display wide interaction capacity in yeast assays. Generally, E function proteins are active when tested in pairwise protein-protein interaction assays (e.g. de Folter *et al.* 2005, Leseberg *et al*. 2008). Arabidopsis SEP1 and SEP3 illustrate an example by providing transcriptional activation for multiple protein complexes (de

Folter *et al*. 2005). Honma and Goto (2001) showed that at least one redundant E function protein is necessary in a complex determining floral organ identity together with A, B, and C function proteins. In contrast to Arabidopsis, where redundancy is extensive among the SEP proteins, Gerbera has specialised E function proteins (GRCD1, GRCD2) involved in determination of stamen and carpel identity together with the C function genes GAGA1 and GAGA2 (Kotilainen *et al*. 2000, Uimari *et al.* 2004). In retrospect, the limited interaction capacity of GRCD1 and GRCD2 nearly exclusively with the C function proteins, reflects their specialized function.

SEP-like proteins are thought to provide the necessary transcriptional activation in tetrameric complexes of MADS domain proteins (Honma & Goto 2001). This phenomenon was carefully studied further, and we discovered similar results for the GAGA1/GRCD2, GAGA2/GRCD2, and GSQUA2/GRCD2 dimers, which reacted under stringent selection conditions. All dimers mentioned above were formed in pairwise assays, but none of them reacted when tested for autoactivation. This raises the question whether the transcriptional activation is the combined function of the interacting proteins, rather than the action of a single protein providing the transcriptional activation signals.

4.7 Gh-SOC1 (III)

Gh-SOC1 was isolated as a full-length cDNA clone from the Gerbera EST collection (Laitinen *et al.* 2005). Gh-SOC1 sequence was highly similar to that of Petunia proteins FBP22, FBP28 and Arabidopsis AGL42 and SOC1. The deduced Gh-SOC1 polypeptide contained a recognizable SOC1 protein motif (Vandenbussche *et al*. 2003b, Nakamura *et al*. 2005) at its C terminal end. In the phylogenetic analysis Gh-SOC1 was closest to Arabidopsis SOC1 and petunia UNS.

Gerbera *Gh-SOC1* overexpression lines did not result in accelerated flowering, which was discovered in many other species ectopically expressing *SOC1*-like genes (Ferrario *et al*. 2004, Vandenbussche *et al*. 2003a). In contrast to many other *SOC1*-like genes, *Gh-SOC1* expression was inflorescence abundant, and no expression was seen invegetative parts. However, overexpression of *Gh-SOC1* interfered with floral development. As a result, the leaf identity expanded to floral organs (III).

4.8 *Gh-SOC1* expression is floral abundant

The expression of *Gh-SOC1* was confined to floral organs in contrast to many closely related genes. Strong expression was seen in involucral bracts, young inflorescences, petals, stamens and carpels. All vegetative parts lacked Gh-SOC1 expression. The timing of *Gh-SOC1* expression was at the later stages of Gerbera petal development (stages 6-11, for stages see Helariutta *et al.* 1993). Interestingly, the expression was stronger in ray than in trans or disc flowers. *In situ* hybrization of young developing gerbera capitula confirmed the expression in petals, ovary area and ovule. *Gh-SOC1* expression was also seen in pappus bristles and ovules, in contrast to the RNA gel blot analysis. *Gh-SOC1* expression was further localized in developing vascular bundles. In disc flowers, the expression of *Gh-SOC1* was shown to be concentrated in the stamens and in the epidermal cell layers of carpel and petals.

Despite being orthologous to Arabidopsis *SOC1* and Petunia *UNS*, the expression of *Gh-SOC1* is only visible in inflorescence and in floral organs. Arabidopsis *SOC1* showed wider range of expression in the leaves, bolting inflorescence, roots and flowers (Borner *et al*. 2000). Also, Petunia *UNS* is expressed in vegetative parts, and only weakly in the inflorescence meristem (Ferrario *et al*. 2004). *Gh-SOC1* showed late petal expression both by RNA gel blot and microarray expression (Laitinen *et al*. 2005). *In situ* hybridization showed clear expression in petals also at earlier development stages. The reason for the lack of the expression at earlier stages in the RNA gel blot analysis was not discovered, and quantitative comparison of the methods employed is difficult.

4.9 Overexpression of *Gh-SOC1* expands vegetative characteristics into Gerbera inflorescence

Three individual lines of overexpression of *Gh-SOC1* showed phenotypic changes. No phenotype was achieved for the downregulation lines. The abaxial side of the ray and trans flower petals featured a greenish tint in contrast to the wild type cultivar ‘Terra Regina’. The adaxial side of the petals had a darker colour than in the non-transgenic gerbera due to doubled anthocyanin content. The disc flowers seemed to have lost some of their pigmentation compared to the wild type disc flowers. In addition, the ray flower petals were shorter, and their surface cell structure was abnormal. By measuring roughly 40 epidermal cells of Gerbera ray flower petals overexpressing *Gh-SOC1*, a difference in cell length was observed: the cells were 8% shortercompared to wild type cells. However, the difference was not statistically significant. The carpel epidermis of overexpression lines contained structures that resembled guard cell stomata. Stomata are typically found in the abaxial surface of the leaves and petals of gerbera, not in the wild type gerbera carpel. No effect on the flowering time was observed.

In Arabidopsis and Petunia, similar phenotypic changes by altered expression of the related genes *SOC1* and *UNS*, respectively, have been reported (Borner et al. 2000, Ferrario et al. 2004). The overexpression of Arabidopsis *SOC1* led to small, green sepaloid petals in Arabidopsis, especially under short day conditions (Borner *et al*. 2000). In Petunia, overexpression of *UNS* gave rise to flowers of greenish hue. Petunia wild type petal epidermal cells are typically ridged, and the flattened epidermal cells discovered in the overexpression line petals were interpreted as a feature of leaves. The overexpression lines also had trichomes on both sides of the petals, in the carpel, and in the ovary. This led the authors to suggest that leaf-like characteristics were the outcome of the ectopic expression of Petunia *UNS*. Ferrario *et al.* (2004) concluded that Petunia flowers lost petal and carpel identity in these lines. In contrast to Arabidopsis *SOC1* and Petunia *UNS*, no effect on the flowering time was observed in the *Gh-SOC1* overexpression lines. Typically, *SOC*-like genes have been reported to accelerate flowering when overexpressed (Borner *et al*. 2000, Ferrario *et al*. 2004, Ma *et al*. 2011), but the only observation of this phenomenon in Gerbera was linked to the overexpression of *GSQUA2* (I). No phenotype was observed in the downregulation lines of *Gh-SOC1*. Both *Gh-SOC2* and *Gh-SOC3* remain interesting candidates to be compared to both Arabidopsis *SOC1* and Petunia *UNS*.

4.10 Gh-SOC1 shows limited interactions with other Gerbera MADS domain proteins

Gh-SOC1 was tested in a pairwise assay with 15 other Gerbera MADS domain proteins. Gh-SOC1 was shown to interact with the euAP1-like proteins GSQUA1 and GSQUA3 (I) and the SEP-like GRCD5 (Yu *et al*. 1999, Laitinen *et al.* 2005, II). GRCD5 and GRCD3 both are linked to petal development in Gerbera (Zhang *et al.* 2017). Gh-SOC1 formed several higher order complexes outside the interactions discovered in the pairwise assays. The higher order complexes were detected among Gh-SOC1, GRCD2/GRCD4 (SEP-like proteins) (Uimari *et al.* 2004, Laitinen *et al.* 2005, II) and GDEF2 (B function protein) (Yu *et al*. 1999). Another SEP-like protein, GRCD3 (Kotilainen *et al*. 2000, Laitinen *et al*. 2005), was bridged by Gh-SOC1 with the C function genes GAGA1 and GAGA2 (Yu *et al.* 1999). The flowering promoting FUL-like GSQUA2 (I) interacted with Gh-SOC1 and the B function genes GGLO1 and GDEF2 (Yu *et al.* 1999). Similar interaction was seen also with GDEF1, which has a function in the stamen development (Broholm *et al*. 2010). Comparable interactions between SOC-like, AP1/FUL-like and B function proteins have not been detected in Arabidopsis, Petunia, or Antirrhinum and the significance of this interaction in Gerbera is unclear. Typically, other model plant B function proteins only interact in higher order complexes as heterodimers. Gerbera B function proteins are active in higher order complexes as single proteins. Based on the PPI data, Gh-SOC1 might have both an early and a late function in Gerbera floral development. In Y3H analysis Gh-SOC1 interacted in different complexes with GSQUA2 and GRCD2 that both have an early function in Gerbera floral meristem development (I, Uimari *et al.*

2004). The interaction with GSQUA2 might also be related to flowering time control. The late function of Gh-SOC1 together with GRCD4 and B function proteins might be related to Gerbera petal development (Yu *et al.* 1999, Broholm *et al*. 2010, Zhang *et al*. 2017).

Arabidopsis SOC1 converges several developmental pathways and has over 20 interacting protein partners in the pairwise assays as shown by Pelaz *et al*. (2001a) and de Folter *et al.* (2005). Gh-SOC1 had a limited interaction capacity in the pairwise assays (II). The detected protein-protein interactions are comparable to those of Arabidopsis SOC1 (de Folter *et al.* 2005). In the higher order complexes, Arabidopsis SOC1 was detected in several complexes (Immink *et al*. 2009). Nine higher order complexes containing Gh-SOC1 were identified, but the interactions were weak compared to other MADS domain proteins tested simultaneously (II). Nevertheless, they can still be significant in terms of floral development.

Immink *et al.* (2012) have studied Arabidopsis *SOC1* further by identifying its upstream and downstream regulators. *SOC1* integrates several flowering time signals (Lee & Lee 2010). The Activity of *SOC1* must be tightly controlled for the formation of perfect flowers (Borner *et al*. 2000). *SOC1* binds to several flowering time regulatory genes that have binding site in *SOC1* (Tao *et al.* 2012) and forms a double negative feedback loop. In addition to flowering time related genes, *SOC1* binds to loci of various miRNA that also control flowering time (Immink *et al.* 2012). *SEP3* and B and C function genes must be suppressed in inflorescence meristem and young floral meristem to avoid premature differentiation. *SOC1, SVP* and *AGL24* are involved in this suppression (Gregis *et al*. 2008, Liu *et al*. 2008). The joint action of these genes creates a time lag in differentiation and allows cells necessary for reproductive organs to be established in the inner whorls. According to Immink *et al.* (2012) SOC1 directly represses B function genes *AP3* and *PI*.

5 CONCLUSIONS AND FUTURE CHALLENGES

Science is a quest to reach better explanations about the world, to test hypotheses and to move forward. Despite of increasing knowledge on floral development in intensively studied model species, like Arabidopsis, Antirrhinum and Petunia, studies on diverse plant taxa are necessary to get a wider perspective on the developmental processes. Over the years, Gerbera has proven to be an excellent model plant for Asteraceae flower development (reviewed in Teeri *et al.* 2006, Elomaa *et al.* 2018). Identification of the molecular building blocks of the complex Gerbera inflorescence is challenging. Gerbera displays traits that are shared with other plant species, but also specific features of its own. Studies on the divergent features provide us means to understand how developmental mechanisms have evolved over the millions of years.

Protein-protein interaction (PPI) studies can reveal differences in behaviour of closely related proteins, which cannot always be detected otherwise. Together with the gene expression data, PPI results can help to develop new hypotheses for gene and protein function. All this adds to our growing knowledge of how proteins act together to exert their function. The most common way of testing protein-protein interactions, the yeast two-hybrid method, can be described as ‘a quick and dirty method’. Yeast is a heterologous organism, and the interactions detected in yeast should be verified by other methods. So far, all the methods developed for this purpose (*in planta)* have their drawbacks. Often, they require advanced technology to detect the interaction.

The scientific methods employed in this study formed tiny steps that helped to solve a small fragment of the puzzle and created new questions. Despite the progress in floral development studies over the last decades, a huge number of unanswered questions remain, and new questions arise as more information is gathered. The amount of raw data has grown as high through-put sequencing methods became available and affordable, and whole plant genomes were sequenced starting with The Arabidopsis Genome Initiative 2000. Also simultaneous analysis of expression profiles of all genes in different environmental conditions, or plant organs, has greatly increased the amount of data available for plant molecular geneticists. The era of functional genetics has changed to the era of proteomics and metabolomics. Bioinformatics provides methods for interpreting these vast interconnected networks. Based on sequence information, computational approaches are developed for predicting e.g. protein-protein interactions. Genome data mining tools are reviewed e.g. in Lee & Kim (2019). The hypotheses generated by computational approaches and protein data have already cut down the number of experiments done in wet lab and released resources for testing only the significant interactions *in situ* and *in planta.* The development of robotic systems allows researchers to test the predicted models in high-throughput systems.

Genome editing by CRISPR/Cas9 method allows very precise manipulation of genetic material and the method is applicable to plants (reviewed in Vats *et al*. 2019). Efficient specific suppression of genes allows gene function to be characterized and multiple genetic loci can be modified simultaneously. In addition to gene suppression, genes can also be introduced to genome by this method. This is of particular interest in crop plant breeding. An additional advantage of CRISPR/Cas 9 method in breeding is lack of foreign genetic material that leads to non-transgenic status and bypasses many legislation issues. Despite the current relative easiness to modify a genome by removing or adding genes, eventually functional studies are still essential to characteririze gene function.

6 ACKNOWLEDGEMENTS

The research for this thesis was carried out during years 2001-2009 in Gerbera Laboratory. Gerbera laboratory was part of Institute of Biotechnology, University of Helsinki until 2002, and later became affiliated with Department of Applied Biology, University of Helsinki (currently Department of Agricultural Sciences, Faculty of Agriculture). Heads of both departments are thanked for providing excellent research facilities. This study was financed by Finnish Academy, Center of Excellence and TEKES, the ESGEMO Research Program.

My supervisor professor Teemu Teeri is thanked for his patience while this work was carried out. He has generously allowed me to pursue my own research interests in addition to the guidelines defined together. He never gave up on my thesis and kept sending kind reminders for ten years. Professor Paula Elomaa is thanked for constantly supplying me and the whole Gerbera group with ideas on how to proceed in research. She has an outstanding ability to inspire people around her and give insightful advice in various research dilemmas.

I thank all my co-authors for their contribution in the publication of my work. In the early stages of my Gerbera research Docent Mika Kotilainen and Dr. Anne Uimari generously shared their extensive knowledge in plant molecular biology and the fascinating world of Gerbera.

Eija Takala I want to thank for her efficient technical assistance and friendship. Marja Huovila and Anu Rokkanen are thanked for their generousness to introduce me to life in laboratory when I arrived in Gerbera lab. Gardener Sanna Peltola is thanked for taking good care of my transgenic Gerbera lines.

I greatly enjoyed collaboration with my former colleagues Drs. Yan Peng Ng and Ursula Malm who were involved in the studies presented here. All the present and former members of Gerbera laboratory are thanked for the nice working atmosphere and all the recreational activities.

My Dutch connection and collaborator Professor Richard Immink, Wageningen University, the Netherlands, is thanked for generously sharing all the information and material. Professor Victor Albert, University of Buffalo, US, wizard of phylogenetics has contributed a lot to my research in the form of phylogenetic analysis.

My parents Ulla and Seppo are thanked for encouraging me and believing in me. This thesis is dedicated to you because you did not have an opportunity to study when you were young.

My rocking husband Ismo, who has offered both mental and financial support and has never demanded me to get 'a real job', I want to thank him for his eternal patience. We have survived working on three tough projects simultaneously (thesis, house and first baby) and are still happily married so I believe we will have many merry moments ahead of us. Besides rocking, you are my rock!

Last and first, I thank my daughter Siiri and my son Iiro. You are the apples of my eye and bring me the greatest joy in my life simply by existing. Genetic recombination produces new phenotypes that are far superior to the ancestral ones.

Vantaa, January 2020
Satu Ruokolainen

7 REFERENCES

Aerts N, de Bruijn S, van Mourik H, Angenent GC, van Dijk ADJ (2018) Comparative analysis of binding patterns of MADS-domain proteins in Arabidopsis thaliana. BMC Plant Biol **18:** 131-018-1348-8

Airoldi CA, Davies B (2012) Gene duplication and the evolution of plant MADS-box transcription factors. J Genet Genomics **39:** 157-165

Almeida AM, Yockteng R, Otoni WC, Specht CD (2015) Positive selection on the K domain of the AGAMOUS protein in the Zingiberales suggests a mechanism for the evolution of androecial morphology. Evodevo **6:**

Alvarez-Buylla E, García-Ponce B, Sánchez Md,la Paz, Espinosa-Soto C, García-Gómez ML, Piñeyro-Nelson A, Garay-Arroyo A (2019) MADS-box genes underground becoming mainstream: plant root developmental mechanisms. New Phytol **0:**

Alvarez-Buylla ER, Pelaz S, Liljegren SJ, Gold SE, Burgeff C, Ditta GS, Ribas de Pouplana L, Martinez-Castilla L, Yanofsky MF (2000) An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. Proc Natl Acad Sci U S A **97:** 5328-5333

Angenent GC, Busscher M, Franken J, Mol JN, van Tunen AJ (1992) Differential expression of two MADS box genes in wild-type and mutant petunia flowers. Plant Cell **4:** 983-993

Angenent GC, Colombo L (1996) Molecular control of ovule development. Trends Plant Sci **1:** 228-232

Angenent GC, Franken J, Busscher M, Colombo L, van Tunen AJ (1993) Petal and stamen formation in petunia is regulated by the homeotic gene fbp1. Plant J **4:** 101-112

Angenent GC, Franken J, Busscher M, van Dijken A, van Went JL, Dons HJ, van Tunen AJ (1995) A novel class of MADS box genes is involved in ovule development in petunia. Plant Cell **7:** 1569-1582

Angenent GC, Franken J, Busscher M, Weiss D, van Tunen AJ (1994) Co-suppression of the petunia homeotic gene fbp2 affects the identity of the generative meristem. Plant J **5:** 33-44

Badouin H, Gouzy J, Grassa CJ, Murat F, Staton SE, Cottret L, Lelandais-Briere C, Owens GL, Carrere S, Mayjonade B, Legrand L, Gill N, Kane NC, Bowers JE, Hubner S, Bellec A, Berard A, Berges H, Blanchet N, Boniface MC, Brunel D, Catrice O, Chaidir N, Claudel C, Donnadieu C, Faraut T, Fievet G, Helmstetter N, King M, Knapp SJ, Lai Z, Le Paslier MC, Lippi Y, Lorenzon L, Mandel JR, Marage G, Marchand G, Marquand E, Bret-Mestries E, Morien E, Nambeesan S, Nguyen T, Pegot-Espagnet P, Pouilly N, Raftis F, Sallet E, Schiex T, Thomas J, Vandecasteele C, Vares D, Vear F, Vautrin S, Crespi M, Mangin B, Burke JM, Salse J, Munos S, Vincourt P, Rieseberg LH, Langlade NB (2017) The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. Nature **546:** 148-152

Balanza V, Martinez-Fernandez I, Ferrandiz C (2014) Sequential action of FRUITFULL as a modulator of the activity of the floral regulators SVP and SOC1. J Exp Bot **65:** 1193-1203

Barker MS, Kane NC, Matvienko M, Kozik A, Michelmore RW, Knapp SJ, Rieseberg LH (2008) Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. Mol Biol Evol **25:** 2445-2455

Becker A, Theissen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol Phylogenet Evol **29:** 464-489

Bemer M, van Mourik H, Muino JM, Ferrandiz C, Kaufmann K, Angenent GC (2017) FRUITFULL controls SAUR10 expression and regulates Arabidopsis growth and architecture. J Exp Bot **68:** 3391-3403

Berbel A, Navarro C, Ferrándiz C, Canas LA, Madueno F, Beltran J- (2001) Analysis of *PEAM4*, the pea *AP1* functional homologue, supports a model for *AP1*-like genes controlling both floral meristem and floral organ identity in different plant species. Plant J **25:** 441-451

Blackman BK, Scascitelli M, Kane NC, Luton HH, Rasmussen DA, Bye RA, Lentz DL, Rieseberg LH (2011) Sunflower domestication alleles support single domestication center in eastern North America. Proc Natl Acad Sci U S A **108:** 14360-14365

Blanc G, Wolfe KH (2004) Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. Plant Cell **16:** 1679-1691

Bontinck M, Van Leene J, Gadeyne A, De Rybel B, Eeckhout D, Nelissen H, De Jaeger G (2018) Recent Trends in Plant Protein Complex Analysis in a Developmental Context. Front Plant Sci **9:** 640

Borner R, Kampmann G, Chandler J, Gleissner R, Wisman E, Apel K, Melzer S (2000) A MADS domain gene involved in the transition to flowering in Arabidopsis. Plant J **24:** 591-599

Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature **422:** 433-438

Bowman JL (1997) Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. Journal of Biosciences **22:** 515

Bowman JL, Alvarez J, Weigel D, Meyerowitz EM, Smyth DR (1993) Control of flower development in Arabidopsis thaliana by APETALA 1 and interacting genes. Development **119:** 721-743

Bowman JL, Drews GN, Meyerowitz EM (1991) Expression of the Arabidopsis floral homeotic gene AGAMOUS is restricted to specific cell types late in flower development. Plant Cell **3:** 749-758

Broholm SK, Pollanen E, Ruokolainen S, Tahtiharju S, Kotilainen M, Albert VA, Elomaa P, Teeri TH (2010) Functional characterization of B class MADS-box transcription factors in Gerbera hybrida. J Exp Bot **61:** 75-85

Broholm SK, Tahtiharju S, Laitinen RA, Albert VA, Teeri TH, Elomaa P (2008) A TCP domain transcription factor controls flower type specification along the radial axis of the Gerbera (Asteraceae) inflorescence. Proc Natl Acad Sci U S A **105:** 9117-9122

Busi MV, Bustamante C, D'Angelo C, Hidalgo-Cuevas M, Boggio SB, Valle EM, Zabaleta E (2003) MADS-box genes expressed during tomato seed and fruit development. Plant Mol Biol **52:** 801-815

Buzgo M, Soltis DE, Soltis PS, Ma H (2004) Towards a comprehensive integration of morphological and genetic studies of floral development. Trends Plant Sci **9:** 164-173

Cartolano M, Castillo R, Efremova N, Kuckenberg M, Zethof J, Gerats T, Schwarz-Sommer Z, Vandenbussche M (2007) A conserved microRNA module exerts homeotic control over Petunia hybrida and Antirrhinum majus floral organ identity. Nat Genet **39:** 901-905

Castelan-Munoz N, Herrera J, Cajero-Sanchez W, Arrizubieta M, Trejo C, Garcia-Ponce B, Sanchez MP, Alvarez-Buylla ER, Garay-Arroyo A (2019) MADS-Box Genes Are Key Components of Genetic Regulatory Networks Involved in Abiotic Stress and Plastic Developmental Responses in Plants. Front Plant Sci **10:** 853

Castillejo C, Romera-Branchat M, Pelaz S (2005) A new role of the Arabidopsis *SEPALLATA3* gene revealed by its constitutive expression. Plant J **43:** 586-596

Causier B, Castillo R, Zhou J, Ingram R, Xue Y, Schwarz-Sommer Z, Davies B (2005) Evolution in action: following function in duplicated floral homeotic genes. Curr Biol **15:** 1508-1512

Chapman MA, Leebens-Mack JH, Burke JM (2008) Positive selection and expression divergence following gene duplication in the sunflower CYCLOIDEA gene family. Mol Biol Evol **25:** 1260-1273

Chapman MA, Tang S, Draeger D, Nambeesan S, Shaffer H, Barb JG, Knapp SJ, Burke JM (2012) Genetic analysis of floral symmetry in Van Gogh's sunflowers reveals independent recruitment of CYCLOIDEA genes in the Asteraceae. PLoS Genet **8:** e1002628

Ciannamea S, Kaufmann K, Frau M, Tonaco IAN, Petersen K, Nielsen KK, Angenent GC, Immink RGH (2006) Protein interactions of MADS box transcription factors involved in flowering in *Lolium perenne*. J Exp Bot **57:** 3419-3431

Coen ES, Meyerowitz EM (1991) The war of the whorls: genetic interactions controlling flower development. Nature **353:** 31-37

Colombo L, Franken J, Koetje E, van Went J, Dons HJ, Angenent GC, van Tunen AJ (1995) The petunia MADS box gene *FBP11* determines ovule identity. Plant Cell **7:** 1859-1868

Colombo L, Franken J, Van der Krol AR, Wittich PE, Dons HJ, Angenent GC (1997) Downregulation of ovule-specific MADS box genes from petunia results in maternally controlled defects in seed development. Plant Cell **9:** 703-715

Cseke LJ, Zheng J, Podila GK (2003) Characterization of PTM5 in aspen trees: a MADS-box gene expressed during woody vascular development. Gene **318:** 55-67

Davies B, Cartolano M, Schwarz-Sommer Z (2006) Flower Development: The *Antirrhinum* perspective. Advances in Botanical Research **44:** 279

Davies B, Motte P, Keck E, Saedler H, Sommer H, Schwarz-Sommer Z (1999) PLENA and FARINELLI: redundancy and regulatory interactions between two Antirrhinum MADS-box factors controlling flower development. EMBO J **18:** 4023-4034

De Bodt S, Raes J, Florquin K, Rombauts S, Rouze P, Theissen G, Van de Peer Y (2003a) Genomewide structural annotation and evolutionary analysis of the type I MADS-box genes in plants. J Mol Evol **56:** 573-586

De Bodt S, Raes J, Van de Peer Y, Theissen G (2003b) And then there were many: MADS goes genomic. Trends Plant Sci **8:** 475-483

de Folter S, Angenent GC (2006) trans meets cis in MADS science. Trends Plant Sci **11:** 224-231

De Folter S, Immink RGH, Kieffer M, Parenicova L, Henz SR, Weigel D, Busscher M, Kooiker M, Colombo L, Kater MM, Davies B, Angenent GC (2005) Comprehensive interaction map of the Arabidopsis MADS box transcription factors. Plant Cell **17:** 1424-1433

Dean EJ, Davis JC, Davis RW, Petrov DA (2008) Pervasive and persistent redundancy among duplicated genes in yeast. PLoS Genet **4:** e1000113

Decroocq V, Zhu X, Kauffman M, Kyozuka J, Peacock WJ, Dennis ES, Llewellyn DJ (1999) A TM3-like MADS-box gene from Eucalyptus expressed in both vegetative and reproductive tissues. Gene **228:** 155-160

DeLuna A, Vetsigian K, Shoresh N, Hegreness M, Colon-Gonzalez M, Chao S, Kishony R (2008) Exposing the fitness contribution of duplicated genes. Nat Genet **40:** 676-681

Ditta G, Pinyopich A, Robles P, Pelaz S, Yanofsky MF (2004) The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. Curr Biol **14:** 1935-1940

Drews GN, Bowman JL, Meyerowitz EM (1991) Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the APETALA2 product. Cell **65:** 991

Edger PP, Pires JC (2009) Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes. Chromosome Res **17:** 699-717

Egea-Cortines M, Saedler H, Sommer H (1999) Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in *Antirrhinum majus*. EMBO J **18:** 5370-5379

Elo A, Lemmetyinen J, Turunen ML, Tikka L, Sopanen T (2001) Three MADS-box genes similar to *APETALA1* and *FRUITFULL* from silver birch (*Betula pendula*). Physiol Plant **112:** 95-103

Elomaa P, Zhao Y, Zhang T (2018) Flower heads in Asteraceae-recruitment of conserved developmental regulators to control the flower-like inflorescence architecture. Hortic Res **5:** 36-018-0056-8. eCollection 2018

Fang SC, Fernandez DE (2002) Effect of regulated overexpression of the MADS domain factor AGL15 on flower senescence and fruit maturation. Plant Physiol **130:** 78-89

Favaro R, Pinyopich A, Battaglia R, Kooiker M, Borghi L, Ditta G, Yanofsky MF, Kater MM, Colombo L (2003) MADS-box protein complexes control carpel and ovule development in Arabidopsis. Plant Cell **15:** 2603-2611

Fernando DD, Zhang S (2006) Constitutive expression of the SAP1 gene from willow (Salix discolor) causes early flowering in Arabidopsis thaliana. Dev Genes Evol **216:** 19-28

Ferrándiz C (2002) Regulation of fruit dehiscence in Arabidopsis. J Exp Bot **53:** 2031-2038

Ferrandiz C, Gu Q, Martienssen R, Yanofsky MF (2000) Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. Development **127:** 725-734

Ferrario S, Busscher J, Franken J, Gerats T, Vandenbussche M, Angenent GC, Immink RGH (2004) Ectopic expression of the petunia MADS box gene *UNSHAVEN* accelerates flowering and confers leaf-like characteristics to floral organs in a dominant-negative manner. Plant Cell **16:** 1490-1505

Ferrario S, Immink RGH, Shchennikova A, Busscher-Lange J, Angenent GC (2003) The MADS box gene *FBP2* is required for *SEPALLATA* function in petunia. Plant Cell **15:** 914-925

Ferrario S, Shchennikova AV, Franken J, Immink RGH, Angenent GC (2006) Control of floral meristem determinacy in petunia by MADS-box transcription factors. Plant Physiol **140:** 890-898

Fornara F, Parenicova L, Falasca G, Pelucchi N, Masiero S, Ciannamea S, Lopez-Dee Z, Altamura MM, Colombo L, Kater MM (2004) Functional characterization of OsMADS18, a member of the AP1/SQUA subfamily of MADS box genes. Plant Physiol **135:** 2207-2219

Freeling M, Thomas BC (2006) Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. Genome Res **16:** 805-814

Friis EM, Pedersen KR, Crane PR (2000) Reproductive structure and organization of basal angiosperms from the Early Cretaceous (Barremian or Aptian) of Western Portugal. International journal of plant sciences **161:** 169-182

Galichet A, Gruissem W (2003) Protein farnesylation in plants--conserved mechanisms but different targets. Curr Opin Plant Biol **6:** 530-535

Goto K, Meyerowitz EM (1994) Function and regulation of the Arabidopsis floral homeotic gene *PISTILLATA*. Genes Dev **8:** 1548-1560

Gramzow L, Ritz MS, Theissen G (2010) On the origin of MADS-domain transcription factors. Trends Genet **26:** 149-153

Gramzow L, Theissen G (2015) Phylogenomics reveals surprising sets of essential and dispensable clades of MIKC(c)-group MADS-box genes in flowering plants. J Exp Zool B Mol Dev Evol **324:** 353-362

Gramzow L, Theissen G (2013) Phylogenomics of MADS-Box Genes in Plants - Two Opposing Life Styles in One Gene Family. Biology (Basel) **2:** 1150-1164

Gramzow L, Theissen G (2010) A hitchhiker's guide to the MADS world of plants. Genome Biol **11:** 214-2010-11-6-214. Epub 2010 Jun 28

Gregis V, Sessa A, Colombo L, Kater MM (2008) AGAMOUS-LIKE24 and SHORT VEGETATIVE PHASE determine floral meristem identity in Arabidopsis. Plant J **56:** 891-902

Gregis V, Sessa A, Colombo L, Kater MM (2006) *AGL24*, *SHORT VEGETATIVE PHASE*, and *APETALA1* redundantly control *AGAMOUS* during early stages of flower development in *Arabidopsis*. Plant Cell **18:** 1373-1382

Gu Q, Ferrandiz C, Yanofsky MF, Martienssen R (1998) The *FRUITFULL* MADS-box gene mediates cell differentiation during Arabidopsis fruit development. Development **125:** 1509-1517

Gustafson-Brown C, Savidge B, Yanofsky MF (1994) Regulation of the arabidopsis floral homeotic gene *APETALA1*. Cell **76:** 131-143

Hancock JF, Cadwallader K, Paterson H, Marshall CJ (1991) A CAAX or a CAAL motif and a second signal are sufficient for plasma membrane targeting of ras proteins. EMBO J **10:** 4033-4039

Harris EM (1995) Inflorescence and floral ontogeny in Asteraceae: a synthesis of historical and current concepts. Botanical Review **91:** 93

Hart JK, Hannapel DJ (2002) *In situ* hybridization of the MADS-box gene *POTM1* during potato floral development. J Exp Bot **53:** 465-471

He CY, Saedler H (2005) Heterotopic expression of MPF2 is the key to the evolution of the Chinese lantern of *Physalis*, a morphological novelty in Solanaceae. Proceedings of the National Academy of Sciences of the United States of America **102:** 5779

He C, Munster T, Saedler H (2004) On the origin of floral morphological novelties. FEBS Lett **567:** 147-151

Heijmans K, Ament K, Rijpkema AS, Zethof J, Wolters-Arts M, Gerats T, Vandenbussche M (2012) Redefining C and D in the petunia ABC. Plant Cell **24:** 2305-2317

Helariutta Y, Elomaa P, Kotilainen M, Seppänen P, Teeri TH (1993) Cloning of cDNA coding for dihydroflavonol-4-reductase (DFR) and characterization of *dfr* expression in the corollas of *Gerbera hybrida* var. Regina (Compositae). Plant Mol Biol **22:** 183-193

Henschel K, Kofuji R, Hasebe M, Saedler H, Munster T, Theissen G (2002) Two ancient classes of MIKC-type MADS-box genes are present in the moss Physcomitrella patens. Mol Biol Evol **19:** 801-814

Heuer S, Hansen S, Bantin J, Brettschneider R, Kranz E, Lorz H, Dresselhaus T (2001) The maize MADS box gene ZmMADS3 affects node number and spikelet development and is co-expressed with ZmMADS1 during flower development, in egg cells, and early embryogenesis. Plant Physiol **127:** 33-45

Hileman LC, Sundström JF, Litt A, Chen M, Shumba T, Irish VF (2006) Molecular and phylogenetic analyses of the MADS-box gene family in tomato. Mol Biol Evol **23:** 2245-2258

Hirakawa H, Sumitomo K, Hisamatsu T, Nagano S, Shirasawa K, Higuchi Y, Kusaba M, Koshioka M, Nakano Y, Yagi M, Yamaguchi H, Taniguchi K, Nakano M, Isobe SN (2019) De novo whole-genome assembly in *Chrysanthemum seticuspe*, a model species of Chrysanthemums, and its application to genetic and gene discovery analysis . DNA Research **26:** 195

Honma T, Goto K (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature **409:** 525-529

Huang H, Tudor M, Weiss CA, Hu Y, Ma H (1995) The *Arabidopsis* MADS-box gene *AGL3* is widely expressed and encodes a sequence-specific DNA-binding protein. Plant Molecular Biology **28:** 549

Huang CH, Zhang C, Liu M, Hu Y, Gao T, Qi J, Ma H (2016) Multiple Polyploidization Events across Asteraceae with Two Nested Events in the Early History Revealed by Nuclear Phylogenomics. Mol Biol Evol **33:** 2820-2835

Huijser P, Klein J, Lonnig WE, Meijer H, Saedler H, Sommer H (1992) Bracteomania, an inflorescence anomaly, is caused by the loss of function of the MADS-box gene *squamosa* in *Antirrhinum majus*. EMBO J **11:** 1239-1249

Immink RG, Pose D, Ferrario S, Ott F, Kaufmann K, Valentim FL, de Folter S, van der Wal F, van Dijk AD, Schmid M, Angenent GC (2012) Characterization of SOC1's central role in flowering by the identification of its upstream and downstream regulators. Plant Physiol **160:** 433-449

Immink RGH, Ferrario S, Busscher-Lange J, Kooiker M, Busscher M, Angenent GC (2003) Analysis of the petunia MADS-box transcription factor family. Molecular genetics and genomics : MGG **268:** 598-606

Immink RGH, Hannapel DJ, Ferrario S, Busscher M, Franken J, Lookeren Campagne MM, Angenent GC (1999) A petunia MADS box gene involved in the transition from vegetative to reproductive development. Development **126:** 5117-5126

Immink R, Tonaco I, de Folter S, Shchennikova A, van Dijk A, Busscher-Lange J, Borst J, Angenent G (2009) SEPALLATA3: the 'glue' for MADS box transcription factor complex formation. Genome Biol **10:** R24

Irish VF, Kramer EM (1998) Genetic and molecular analysis of angiosperm flower development. Adv Bot Res 197-230

Irish VF, Sussex IM (1990) Function of the *apetala-1* gene during Arabidopsis floral development. Plant Cell **2:** 741-753

Ito T, Ng K-, Lim T-, Yu H, Meyerowitz EM (2007) The Homeotic Protein AGAMOUS Controls Late Stamen Development by Regulating a Jasmonate Biosynthetic Gene in *Arabidopsis*. Plant Cell **19:** 3516-3529

Jaakola L, Poole M, Jones MO, Kamarainen-Karppinen T, Koskimaki JJ, Hohtola A, Haggman H, Fraser PD, Manning K, King GJ, Thomson H, Seymour GB (2010) A SQUAMOSA MADS box gene involved in the regulation of anthocyanin accumulation in bilberry fruits. Plant Physiol **153:** 1619-1629

Jack T, Brockman LL, Meyerowitz EM (1992) The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. Cell **68:** 683-697

Jang S, An K, Lee S, An G (2002) Characterization of tobacco MADS-box genes involved in floral initiation. Plant Cell Physiol **43:** 230-238

Jeffrey C (2007) Compositae: Introduction with key to tribes. In K Kubitzki, ed, The families and genera of vascular plants. Springer Verlag, Berlin, pp 61

Jeon J-, Lee S, Jung K-, Jun S-, Jeong D-, Lee J, Kim C, Jang S, Lee S, Yang K, Nam J, An K, Han M-, Sung R-, Choi H-, Yu J-, Choi J-, Cho S-, Cha S-, Kim S-, An G (2000) **T-DNA insertional mutagenesis for functional genomics in rice**. The Plant Journal **22:** 561

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, de-Pamphilis CW (2011) Ancestral polyploidy in seed plants and angiosperms. Nature **473:** 97-100

Jofuku KD, den Boer BG, Van Montagu M, Okamuro JK (1994) Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. Plant Cell **6:** 1211-1225

Kafri R, Springer M, Pilpel Y (2009) Genetic redundancy: new tricks for old genes. Cell **136:** 389-392

Kanno A, Saeki H, Kameya T, Saedler H, Theissen G (2003) Heterotopic expression of class B floral homeotic genes supports a modified ABC model for tulip (Tulipa gesneriana). Plant Molecular Biology **52:** 831

Kapoor M, Tsuda S, Tanaka Y, Mayama T, Okuyama Y, Tsuchimoto S, Takatsuji H (2002) Role of petunia pMADS3 in determination of floral organ and meristem identity, as revealed by its loss of function. Plant J **32:** 115-127

Kater MM, Colombo L, Franken J, Busscher M, Masiero S, Van Lookeren Campagne MM, Angenent GC (1998) Multiple AGAMOUS homologs from cucumber and petunia differ in their ability to induce reproductive organ fate. Plant Cell **10:** 171-182

Kaufmann K, Melzer R, Theißen G (2005) MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. Gene **347:** 183-198

Kaufmann K, Muino JM, Jauregui R, Airoldi CA, Smaczniak C, Krajewski P, Angenent GC (2009) Target genes of the MADS transcription factor SEPALLATA3: integration of developmental and hormonal pathways in the Arabidopsis flower. PLoS Biol **7:** e1000090

Keck E, McSteen P, Carpenter R, Coen E (2003) Separation of genetic functions controlling organ identity in flowers. EMBO J **22:** 1058-1066

Kempin SA, Savidge B, Yanofsky MF (1995) Molecular basis of the cauliflower phenotype in Arabidopsis. Science **267:** 522-525

Kim S, Koh J, Yoo M-, Kong H, Hu Y, Ma H, Soltis PS, Soltis DE (2005) Expression of floral MADS-box genes in basal angiosperms: implications on evolution of floral regulators and the perianth. The Plant Journal **43:** 724

Kim S, Yoo MJ, Albert VA, Farris JS, Soltis PS, Soltis DE (2004) Phylogeny and diversification of B-function MADS-box genes in angiosperms: evolutionary and functional implications of a 260-million-year-old duplication. Am J Bot **91:** 2102-2118

Kofuji R, Sumikawa N, Yamasaki M, Kondo K, Ueda K, Ito M, Hasebe M (2003) Evolution and divergence of the MADS-box gene family based on genome-wide expression analyses. Mol Biol Evol **20:** 1963-1977

Kohler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U (2003) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. Genes Dev **17:** 1540-1553

Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV (2002) Selection in the evolution of gene duplications. Genome biology **3:** research0008.1

Kotilainen M, Elomaa P, Uimari A, Albert VA, Yu D, Teeri TH (2000) *GRCD1*, an *AGL2*-like MADS box gene, participates in the C function during stamen development in *Gerbera hybrida*. Plant Cell **12:** 1893-1902

Kramer EM, Di Stilio VS, Schluter PM (2003) Complex patterns of gene duplication in the *APETALA3* and *PISTILLATA* lineages of the Ranunculaceae. Journal of Plant Sciences **164:** 1

Kramer EM, Dorit RL, Irish VF (1998) Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the *APETALA3* an *PISTILLATA* MADS-box gene lineages. [Erratum: Feb 1999, v. 151 (2), p. 915.]. Genetics **149:** 765-783

Kramer EM, Irish VF (2000) Evolution of petal and stamen developmental programs: Evidence from comparative studies of the lower eudicots and basal angiosperms. International journal of plant sciences **161:** 29

Kramer EM, Irish VF (1999) Evolution of genetic mechanisms controlling petal development. Nature **399:** 144-148

Kramer EM, Jaramillo MA, Di Stilio VS (2004) Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms. Genetics **166:** 1011-1023

Krizek BA, Fletcher JC (2005) Molecular mechanisms of flower development: an armchair guide. Nat Rev Genet **6:** 688-698

Kunst L, Klenz JE, Martinez-Zapater J, Haughn GW (1989) AP2 Gene Determines the Identity of Perianth Organs in Flowers of Arabidopsis thaliana. Plant Cell **1:** 1195-1208

Kyozuka J, Harcourt R, Peacock WJ, Dennis ES (1997) Eucalyptus has functional equivalents of the Arabidopsis *AP1* gene. Plant Mol Biol **35:** 573-584

Laitinen RA, Broholm S, Albert VA, Teeri TH, Elomaa P (2006) Patterns of MADS-box gene expression mark flower-type development in *Gerbera hybrida* (Asteraceae). BMC Plant Biol **6:** 11

Laitinen RA, Immanen J, Auvinen P, Rudd S, Alatalo E, Paulin L, Ainasoja M, Kotilainen M, Koskela S, Teeri TH, Elomaa P (2005) Analysis of the floral transcriptome uncovers new regulators of organ determination and gene families related to flower organ differentiation in *Gerbera hybrida* (Asteraceae). Genome Res **15:** 475-486

Lamb RS, Irish VF (2003) Functional divergence within the APETALA3/PISTILLATA floral homeotic gene lineages. Proc Natl Acad Sci U S A **100:** 6558-6563

Lee J, Lee I (2010) Regulation and function of SOC1, a flowering pathway integrator. Journal of experimental botany **61:** 2247

Lee GW, Kim S (2008) Genome data mining for everyone. BMB Rep **41:** 757-764

Lee H, Suh SS, Park E, Cho E, Ahn JH, Kim SG, Lee JS, Kwon YM, Lee I (2000) The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis. Genes Dev **14:** 2366-2376

Lenhard M, Bohnert A, Jurgens G, Laux T (2001) Termination of stem cell maintenance in *Arabidopsis* floral meristems by interactions between *WUSCHEL* and *AGAMOUS*. Cell **105:** 805-814

Leseberg CH, Eissler CL, Wang X, Johns MA, Duvall MR, Mao L (2008) Interaction study of MADS-domain proteins in tomato. J Exp Bot **59:** 2253-2265

Liljegren SJ, Ferrandiz C, Alvarez-Buylla ER, Pelaz S, Yalovsky S (1998) MADS-box genes involved in fruit dehiscence. Flowering Newsletter **25:** 9

Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF (2000) SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature **404:** 766-770

Liljegren SJ, Roeder AH, Kempin SA, Gremski K, Ostergaard L, Guimil S, Reyes DK, Yanofsky MF (2004) Control of fruit patterning in Arabidopsis by INDEHISCENT. Cell **116:** 843-853

Litt A (2007) An evaluation of A-function: Evidence from the *APETALA1* and *APETALA2* gene lineages. Int J Plant Sci **168:** 73-91

Litt A, Irish VF (2003) Duplication and diversification in the *APETALA1*/*FRUITFULL* floral homeotic gene lineage: implications for the evolution of floral development. Genetics **165:** 821-833

Liu C, Chen H, Er HL, Soo HM, Kumar PP, Han JH, Liou YC, Yu H (2008) Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis.
. Development **135:** 1481

Liu C, Xi W, Shen L, Tan C, Yu H (2009) Regulation of floral patterning by flowering time genes. Dev Cell **16:** 711-722

Liu C, Zhou J, Bracha-Drori K, Yalovsky S, Ito T, Yu H (2007) Specification of Arabidopsis floral meristem identity by repression of flowering time genes. Development **134:** 1901-1910

Lohmann JU, Hong RL, Hobe M, Busch MA, Parcy F, Simon R, Weigel D (2001) A molecular link between stem cell regulation and floral patterning in Arabidopsis. Cell **105:** 793-803

Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science **290:** 1151-1155

Ma G, Ning G, Zhang W, Zhan J, Lv H, Bao M (2011) Overexpression of *Petunia SOC1*-like gene *FBP21* in tobacco promotes flowering without decreasing flower or fruit quantity. Plant Molecular Biology Reporter **29:** 573-581

Ma H, Yanofsky MF, Meyerowitz EM (1991) *AGL1-AGL6*, an Arabidopsis gene family with similarity to floral homeotic and transcription factor genes. Genes Dev **5:** 484-495

Ma L, Sun N, Liu X, Jiao Y, Zhao H, Deng XW (2005) Organ-Specific Expression of Arabidopsis Genome during Development. Plant Physiol **138:** 80-91

Maere S, De Bodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, Van de Peer Y (2005) Modeling gene and genome duplications in eukaryotes. Proc Natl Acad Sci U S A **102:** 5454-5459

Maes T, Van de Steene N, Zethof J, Karimi M, D'Hauw M, Mares G, Van Montagu M, Gerats T (2001) Petunia Ap2-like genes and their role in flower and seed development. Plant Cell **13:** 229-244

Mandel MA, Gustafson-Brown C, Savidge B, Yanofsky MF (1992) Molecular characterization of the Arabidopsis floral homeotic gene *APETALA1*. Nature **360:** 273-277

Mandel MA, Yanofsky MF (1995b) A gene triggering flower formation in Arabidopsis. Nature **377:** 522-524

Mandel MA, Yanofsky MF (1995a) The Arabidopsis *AGL8* MADS box gene is expressed in inflorescence meristems and is negatively regulated by *APETALA1*. Plant Cell **7:** 1763-1771

Masiero S, Colombo L, Grini PE, Schnittger A, Kater MM (2011) The emerging importance of type I MADS box transcription factors for plant reproduction. Plant Cell **23:** 865-872

Melzer R, Verelst W, Theissen G (2009) The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in 'floral quartet'-like complexes in vitro. Nucleic Acids Res **37:** 144-157

Melzer S, Lens F, Gennen J, Vanneste S, Rohde A, Beeckman T (2008) Flowering-time genes modulate meristem determinacy and growth form in Arabidopsis thaliana. Nat Genet **40:** 1489-1492

Mena M, Ambrose BA, Meeley RB, Briggs SP, Yanofsky MF, Schmidt RJ (1996) Diversification of C-function activity in maize flower development. Science **274:** 1537-1540

Meyerowitz EM, Bowman JL, Brockman LL, Drews GN, Jack T, Sieburth LE, Weigel D (1991) A genetic and molecular model for flower development in *Arabidopsis thaliana*. Dev Suppl **1:** 157-167

Moon Y, Kang H, Jung J, Jeon J, Sung S, An G (1999) Determination of the Motif Responsible for Interaction between the Rice APETALA1/AGAMOUS-LIKE9 Family Proteins Using a Yeast Two-Hybrid System. Plant Physiol **120:** 1193-1204

Moore RC, Grant SR, Purugganan MD (2005) Molecular population genetics of redundant floral-regulatory genes in Arabidopsis thaliana. Molecular biology and evolution **22:** 91-103

Morel P, Chambrier P, Boltz V, Chamot S, Rozier F, Rodrigues Bento S, Trehin C, Monniaux M, Zethof J, Vandenbussche M (2019) Divergent Functional Diversification Patterns in the SEP/AGL6/AP1 MADS-Box Transcription Factor Superclade. Plant Cell **31:** 3033-3056

Morel P, Heijmans K, Rozier F, Zethof J, Chamot S, Rodrigues Bento S, Vialette- Guiraud A, Chambrier P, Trehin C, Vandenbussche M (2017) Divergence of the Floral A-Function between an Asterid and a Rosid Species. THE PLANT CELL **29:** 1605-1621

Muller BM, Saedler H, Zachgo S (2001) The MADS-box gene DEFH28 from Antirrhinum is involved in the regulation of floral meristem identity and fruit development. Plant J **28:** 169-179

Münster T, Faigl W, Saedler H, Theissen G (2002) Evolutionary Aspects of MADS-box Genes in the Eusporangiate Fern Ophioglossum. Plant Biology **4:** 474-483

Nakamura T, Song IJ, Fukuda T, Yokoyama J, Maki M, Ochiai T, Kameya T, Kanno A (2005) Characterization of TrcMADS1 gene of Trillium camtschatcense (Trilliaceae) reveals functional evolution of the SOC1/TM3-like gene family. J Plant Res **118:** 229-234

Nam J, dePamphilis CW, Ma H, Nei M (2003) Antiquity and evolution of the MADS-box gene family controlling flower development in plants. Mol Biol Evol **20:** 1435-1447

Nam J, Kim J, Lee S, An G, Ma H, Nei M (2004) Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms. Proc Natl Acad Sci U S A **101:** 1910-1915

Norman C, Runswick M, Pollock R, Treisman R (1988) Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response element. Cell **55:** 989-1003

Nowak MA, Boerlijst MC, Cooke J, Smith JM (1997) Evolution of genetic redundancy. Nature **388:** 167-171

Ó'Maoiléidigh DS, Wuest SE, Rae L, Raganelli A, Ryan PT, Kwaśniewska K, Das P, Lohan AJ, Loftus B, Graciet E, Wellmer F (2013) Control of Reproductive Floral Organ Identity Specification in *Arabidopsis* by the C Function Regulator AGAMOUS. THE PLANT CELL **25:** 2482

Pabon-Mora N, Ambrose BA, Litt A (2012) Poppy APETALA1/FRUITFULL orthologs control flowering time, branching, perianth identity, and fruit development. Plant Physiol **158:** 1685-1704

Panchy N, Lehti-Shiu M, Shiu S (2016) Evolution of Gene Duplication in Plants. Plant Physiology **171:** 2294-2316

Parenicová L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B, Angenent GC, Colombo L (2003) Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. Plant Cell **15:** 1538-1551

Passmore S, Elble R, Tye BK (1989) A protein involved in minichromosome maintenance in yeast binds a transcriptional enhancer conserved in eukaryotes. Genes Dev **3:** 921-935

Passmore S, Maine GT, Elble R, Christ C, Tye BK (1988) Saccharomyces cerevisiae protein involved in plasmid maintenance is necessary for mating of MAT alpha cells. J Mol Biol **204:** 593-606

Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF (2000) B and C floral organ identity functions require *SEPALLATA* MADS-box genes. Nature **405:** 200-203

Pelaz S, Gustafson-Brown C, Kohalmi SE, Crosby WL, Yanofsky MF (2001 A) *APETALA1* and *SEPALLATA3* interact to promote flower development. Plant J **26:** 385-394

Pelaz S, Tapia-Lopez R, Alvarez-Buylla ER, Yanofsky MF (2001 B) Conversion of leaves into petals in Arabidopsis. Curr Biol **11:** 182-184

Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature **424:** 85-88

Pollock R, Treisman R (1990) **A sensitive method for the determination of protein-DNA binding specificities**. Nucleic Acids Research **18:** 6197

Preston JC, Hileman LC (2010) SQUAMOSA-PROMOTER BINDING PROTEIN 1 initiates flowering in Antirrhinum majus through the activation of meristem identity genes. Plant J **62:** 704-712

Purugganan MD (1997) The MADS-box floral homeotic gene lineages predate the origin of seed plants: phylogenetic and molecular clock estimates. J Mol Evol **45:** 392-396

Purugganan MD, Rounsley SD, Schmidt RJ, Yanofsky MF (1995) Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. Genetics **140:** 345-356

Ratcliffe OJ, Kumimoto RW, Wong BJ, Riechmann JL (2003) Analysis of the Arabidopsis MADS AFFECTING FLOWERING gene family: MAF2 prevents vernalization by short periods of cold. Plant Cell **15:** 1159-1169

Ratcliffe OJ, Nadzan GC, Reuber TL, Riechmann JL (2001) Regulation of flowering in Arabidopsis by an *FLC* homologue. Plant Physiol **126:** 122-132

Riechmann JL, Krizek BA, Meyerowitz EM (1996 A) Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. Proc Natl Acad Sci U S A **93:** 4793-4798

Riechmann JL, Meyerowitz EM (1997) Determination of floral organ identity by Arabidopsis MADS domain homeotic proteins AP1, AP3, PI, and AG is independent of their DNA-binding specificity. Mol Biol Cell **8:** 1243-1259

Riechmann JL, Wang M, Meyerowitz EM (1996 B) DNA-binding properties of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. Nucleic Acids Res **24:** 3134-3141

Rijpkema AS, Royaert S, Zethof J, Van Der Weerden G, Gerats T, Vandenbussche M (2006) Analysis of the Petunia *TM6* MADS box gene reveals functional divergence within the *DEF*/*AP3* lineage. Plant Cell **18:** 1819-1832

Rijpkema AS, Zethof J, Gerats T, Vandenbussche M (2009) The petunia AGL6 gene has a SEPALLATA-like function in floral patterning. Plant J **60:** 1-9

Ronse De Craene , L. P., Soltis PS, Soltis DE (2003) Evolution of floral structures in basal angiosperms. International journal of plant sciences **164:** 329-363

Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G (2000) Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. Science **288:** 1613-1616

Scannell DR, Byrne KP, Gordon JL, Wong S, Wolfe KH (2006) Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. Nature **440:** 341-345

Schwarz-Sommer Z, Davies B, Hudson A (2003) An everlasting pioneer: the story of Antirrhinum research. Nat Rev Genet **4:** 657-666

Schwarz-Sommer Z, Hue I, Huijser P, Flor PJ, Hansen R, Tetens F, Lonnig WE, Saedler H, Sommer H (1992) Characterization of the Antirrhinum floral homeotic MADS-box gene *deficiens*: evidence for DNA binding and autoregulation of its persistent expression throughout flower development. EMBO J **11:** 251-263

Shchennikova AV, Shulga OA, Immink R, Skryabin KG, Angenent GC (2004) Identification and characterization of four chrysanthemum MADS-box genes, belonging to the *APETALA1*/*FRUITFULL* and *SEPALLATA3* subfamilies. Plant Physiol **134:** 1632-1641

Shulga OA, Shchennikova AV, Angenent GC, Skryabin KG (2008) MADS-box genes controlling inflorescence morphogenesis in sunflower. Russ J Dev Bio **39:** 2-5

Skipper M, Pedersen KB, Johansen LB, Frederiksen S, Irish VF, Johansen BB (2005) Identification and quantification of expression levels of three *FRUITFULL*-like MADS-box genes from the orchid *Dendrobium thyrsiflorum* (Reichb. f.). Plant Science **169:** 579-586

Smaczniak C, Immink RG, Muino JM, Blanvillain R, Busscher M, Busscher-Lange J, Dinh QD, Liu S, Westphal AH, Boeren S, Parcy F, Xu L, Carles CC, Angenent GC, Kaufmann K (2012) Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. Proc Natl Acad Sci U S A **109:** 1560-1565

Smykal P, Gennen J, De Bodt S, Ranganath V, Melzer S (2007) Flowering of strict photoperiodic Nicotiana varieties in non-inductive conditions by transgenic approaches. Plant Mol Biol **65:** 233-242

Soltis DE, Soltis PS, Chase MW, Endress P (2005) Phylogeny and evolution of angiosperms. Sinauer Associates, Sunderland, MA

Soltis DE, Chanderbali AS, Kim S, Buzgo M, Soltis PS (2007) The ABC model and its applicability to basal angiosperms. Ann Bot **100:** 155-163

Sommer H, Beltran JP, Huijser P, Pape H, Lonnig WE, Saedler H, Schwarz-Sommer Z (1990) *Deficiens*, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. EMBO J **9:** 605-613

Sreekantan L, Clemens J, McKenzie MJ, Lenton JR, Croker SJ, Jameson PE (2004) Flowering genes in *Metrosideros* fit a broad herbaceous model encompassing *Arabidopsis* and *Antirrhinum*. Physiol Plant **121:** 163-173

Sridhar VV, Surendrarao A, Liu Z (2006) *APETALA1* and *SEPALLATA3* interact with *SEUSS* to mediate transcription repression during flower development. Development **133:** 3159-3166

Sung SK, Moon YH, Chung JE, Lee SY, Park HG, An G (2001) Characterization of MADS box genes from hot pepper. Mol Cells **11:** 352-359

Sung SK, Yu GH, An G (1999) Characterization of *MdMADS2*, a member of the *SQUAMOSA* subfamily of genes, in apple. Plant Physiol **120:** 969-978

Svensson ME, Johannesson H, Engstrom P (2000) The LAMB1 gene from the clubmoss, Lycopodium annotinum, is a divergent MADS-box gene, expressed specifically in sporogenic structures. Gene **253:** 31-43

Tadege M, Sheldon CC, Helliwell CA, Upadhyaya NM, Dennis ES, Peacock WJ (2003) Reciprocal control of flowering time by OsSOC1 in transgenic Arabidopsis and by FLC in transgenic rice. Plant Biotechnol J **1:** 361-369

Tahtiharju S, Rijpkema AS, Vetterli A, Albert VA, Teeri TH, Elomaa P (2012) Evolution and diversification of the CYC/TB1 gene family in Asteraceae--a comparative study in Gerbera (Mutisieae) and sunflower (Heliantheae). Mol Biol Evol **29:** 1155-1166

Tandre K, Albert VA, Sundas A, Engstrom P (1995) Conifer homologues to genes that control floral development in angiosperms. Plant Mol Biol **27:** 69-78

Tao Z, Shen L, Liu C, Liu L, Yan Y, Yu H (2012) Genome-wide identification of SOC1 and SVP targets during the floral transition in Arabidopsis. Plant J **70:** 549-561

Teeri TH, Kotilainen M, Uimari A, Ruokolainen S, Ng YP, Malm U, Pöllänen E, Broholm S, Laitinen R, Elomaa P, Albert VA (2006) Floral developmental genetics of *Gerbera* (Asteraceae). Adv Bot Res **44:** 323-351

The Arabidopsis Genome Initiative (2000) **Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana***. Nature **408:** 796

Theissen G, Gramzow L (2016) Structure and evolution of plant MADS-domain transcription factors. In DH Gonzalez, ed, *Plant Transcription Factors: Evolutionary, Structural and Functional Aspects*. Elsevier, Philadelphia, pp 127

Theissen G (2001) Development of floral organ identity: stories from the MADS house. Curr Opin Plant Biol **4:** 75-85

Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, Winter KU, Saedler H (2000) A short history of MADS-box genes in plants. Plant Mol Biol **42:** 115-149

Theissen G, Kim JT, Saedler H (1996) Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. J Mol Evol **43:** 484-516

Theissen G, Melzer R (2007) Molecular mechanisms underlying origin and diversification of the angiosperm flower. Ann Bot **100:** 603-619

Theissen G, Melzer R, Rumpler F (2016) MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. Development **143:** 3259-3271

Theissen G, Saedler H (2001) Plant biology. Floral quartets. Nature **409:** 469-471

Tischler J, Lehner B, Chen N, Fraser AG (2006) Combinatorial RNA interference in Caenorhabditis elegans reveals that redundancy between gene duplicates can be maintained for more than 80 million years of evolution. Genome Biol **7:** R69-2006-7-8-R69. Epub 2006 Aug 2

Trevaskis B, Bagnall DJ, Ellis MH, Peacock WJ, Dennis ES (2003) MADS box genes control vernalization-induced flowering in cereals. Proc Natl Acad Sci U S A **100:** 13099-13104

Tröbner W, Ramirez L, Motte P, Hue I, Huijser P, Lonnig WE, Saedler H, Sommer H, Schwarz-Sommer Z (1992) *GLOBOSA*: a homeotic gene which interacts with *DEFICIENS* in the control of Antirrhinum floral organogenesis. EMBO J **11:** 4693-4704

Tsuchimoto S, van der Krol AR, Chua NH (1993) Ectopic expression of pMADS3 in transgenic petunia phenocopies the petunia blind mutant. Plant Cell **5:** 843-853

Uimari A, Kotilainen M, Elomaa P, Yu D, Albert VA, Teeri TH (2004) Integration of reproductive meristem fates by a *SEPALLATA*-like MADS-box gene. Proc Natl Acad Sci U S A **101:** 15817-15822

Vallade J, Maizonnier D, Cornu A (1987) La morphogenese florale chez le petunia. Analyze d'un mutant à corolle staminée. Canadian Journal of Botany **65:** 761

Van de Peer Y, Maere S, Meyer A (2009) The evolutionary significance of ancient genome duplications. Nat Rev Genet **10:** 725-732

van der Krol AR, Brunelle A, Tsuchimoto S, Chua NH (1993) Functional analysis of petunia floral homeotic MADS box gene pMADS1. Genes Dev **7:** 1214-1228

van Tunen AJ, Eikelboom W, Angenent GC (1993) Floral organosis in *Tulipa*. Flowering Newsletter **16:** 33

Vandenbussche M, Zethof J, Royaert S, Weterings K, Gerats T (2004) The duplicated B-class heterodimer model: Whorl-specific effects and complex genetic interactions in *Petunia hybrida* flower development. Plant Cell **16:** 741-754

Vandenbussche M, Zethof J, Souer E, Koes R, Tornielli GB, Pezzotti M, Ferrario S, Angenent GC, Gerats T (2003a) Toward the analysis of the petunia MADS box gene family by reverse and forward transposon insertion mutagenesis approaches: B, C, and D floral organ identity functions require SEPALLATA-like MADS box genes in petunia. Plant Cell **15:** 2680-2693

Vandenbussche M, Theissen G, Van de Peer Y, Gerats T (2003b) Structural diversification and neo-functionalization during floral MADS-box gene evolution by C-terminal frameshift mutations. Nucl Acids Res **31:** 4401-4409

Vats S, Kumawat S, Kumar V, Patil GB, Joshi T, Sonah H, Sharma TR, Deshmukh R (2019) Genome Editing in Plants: Exploration of Technological Advancements and Challenges. Cells **8:** 10.3390/cells8111386

von Goethe JW (1790) Versuch die metamorphose der pflanzen zu erklären. C.W. Ettinger, Gotha, Germany

Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W, Giovannoni J (2002) A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. Science **296:** 343-346

Wagner A (1999) Redundant gene functions and natural selection. Journal of Evolutionary Biology **12:** 1

Wagner A (2008) Gene duplications, robustness and evolutionary innovations. Bioessays **30:** 367-373

Walden AR, Wang DY, Walter C, Gardner RC (1998) A large family of TM3 MADS-box cDNAs in Pinus radiata includes two members with deletions of the conserved K domain. Plant Science **138:** 167-176

Wang H, Caruso LV, Downie AB, Perry SE (2004) The Embryo MADS Domain Protein AGAMOUS-Like 15 Directly Regulates Expression of a Gene Encoding an Enzyme Involved in Gibberellin Metabolism. Plant Cell **16:** 1206-1219

Watson JM, Brill EM (2004) Eucalyptus grandis has at least two functional SOC1-like floral activator genes. Functional **31:** 31. 10.1071/FP03181

Whipple CJ, Ciceri P, Padilla CM, Ambrose BA, Bandong SL, Schmidt RJ (2004) Conservation of B-class floral homeotic gene function between maize and Arabidopsis. Development **131:** 6083-6091

Winter KU, Becker A, Munster T, Kim JT, Saedler H, Theissen G (1999) MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. Proc Natl Acad Sci U S A **96:** 7342-7347

Wollmann H, Mica E, Todesco M, Long JA, Weigel D (2010) On reconciling the interactions between APETALA2, miR172 and AGAMOUS with the ABC model of flower development. Development **137:** 3633-3642

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Rockman MV, Romano LA (2003) The evolution of transcriptional regulation in eukaryotes. Molecular biology and evolution **9:** 1377

Wu YH, Zhang JS, Zheng Z, Xue S, Li Y (2000) Molecular cloning and characterization of two tobacco MADS-box genes. Sexual plant reproduction **13:** 163-169

Wynne J, Treisman R (1992) SRF and MCM1 have related but distinct DNA binding specificities. Nucleic Acids Research **20:** 3297

Yalovsky S, Kulukian A, Rodriguez-Concepcion M, Young CA, Gruissem W (2000b) Functional requirement of plant farnesyltransferase during development in Arabidopsis. Plant Cell **12:** 1267-1278

Yalovsky S, Rodriguez-Concepcion M, Bracha K, Toledo-Ortiz G, Gruissem W (2000a) Prenylation of the floral transcription factor APETALA1 modulates its function. Plant Cell **12:** 1257-1266

Yan W, Chen D, Kaufmann K (2016) Molecular mechanisms of floral organ specification by MADS domain proteins. Curr Opin Plant Biol **29:** 154-162

Yang Y, Fanning L, Jack T (2003) The K domain mediates heterodimerization of the Arabidopsis floral organ identity proteins, APETALA3 and PISTILLATA. Plant J **33:** 47-59

Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM (1990) The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. Nature **346:** 35-39

Yu DY, Kotilainen M, Pöllänen E, Mehto M, Elomaa P, Helariutta Y, Albert VA, Teeri TH (1999) Organ identity genes and modified patterns of flower development in *Gerbera hybrida* (Asteraceae). Plant J **17:** 51-62

Zahn LM, Kong H, Leebens-Mack JH, Kim S, Soltis PS, Landherr LL, Soltis DE, Depamphilis CW, Ma H (2005) The evolution of the SEPALLATA subfamily of MADS-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history. Genetics **169:** 2209-2223

Zhang H, Forde BG (1998) An Arabidopsis MADS box gene that controls nutrient-induced changes in root architecture. Science **279:** 407-409

Zhang T, Zhao Y, Juntheikki I, Mouhu K, Broholm SK, Rijpkema AS, Kins L, Lan T, Albert VA, Teeri TH, Elomaa P (2017) Dissecting functions of SEPALLATA-like MADS box genes in patterning of the pseudanthial inflorescence of Gerbera hybrida. New Phytol **216:** 939-954

ISBN 978-951-51-6687-6 (PRINT)
ISBN 978-951-51-6688-3 (ONLINE)
ISSN 2342-5423 (PRINT)
ISSN 2342-5431 (ONLINE)
http://ethesis.helsinki.fi

HELSINKI 2020